A composite step method for equality constrained optimization on manifolds

Julian Ortiz & Anton Schiela

March 1, 2019

Abstract

We present a composite step method, designed for equality constrained optimization on differentiable manifolds. The use of retractions allows us to pullback the involved mappings to linear spaces and use tools such as cubic regularization of the objective function and affine covariant damped Newton method for feasibility. We show fast local convergence when different chart retractions are considered. We test our method on equilibrium problems in finite elasticity where the stable equilibrium position of an inextensible transversely isotropic elastic rod under dead load is searched.

AMS MSC 2000: 49M37, 90C55, 90C06

Keywords: composite step methods, retractions, optimization on manifolds

1 Introduction

In an important variety of fields, optimization problems benefit from a formulation on nonlinear manifolds. Problems in numerical linear algebra like invariant subspace computations, or low rank approximation problems can be tackled using this approach, these problems are the focus of [AMS09]. Nonlinear partial differential equations where the configuration space is given by maps where the domain and target are nonlinear manifolds are found im many applications. Examples are Cosserat materials [BS89] where configurations are maps into the space $\mathbb{R}^3 \times SO(3)$ which are particularly relevant for shell and rod mechanics. Liquid crystal physics [Pro95] where molecules are described as little rod- or plate-like objects; in a PDE setting a liquid Crystal configuration is a field with values in the unit sphere, or, depending on the symmetry of the molecules, in the projective plane or the special orthogonal group. Various numerical approaches to simulate liquid crystals and related problems from micro-magnetics can be found in the literature [Alo97, AKT12, BP07, KVBP$^+$14, LL89].

Numerical computations with shapes, such as shape analysis [MB11, RW12] and shape optimization [Sch14] are done, using the inherent structure of the space of shapes. This structure originates from the fact that deformations, i.e., diffeomorphisms form a Lie group, rather than a vector space. Similar insights have been succesfully exploited in the analysis of finite strain elasticity and elastoplasticity [Bal02, Mie02]. Further applications of fields with nonlinear codomain are models of topological solitions [MS04], image processing [TSC00], and the treatment of diffusion-tensor imaging [PFA06]. Mathematical literature can be found in [SS00] on geometric wave maps, or [EL78] on harmonic maps.

Unconstrained optimization on manifolds is by now well established, as can be seen in [AMS09, Lue72, TSC00], where the theory of optimization is covered. Many things run in parallel to algorithmic approaches on linear spaces. In particular, local (usually quadratic) models are minimized at the current iterate, giving rise to the construction of the next step. The main difference between optimization algorithms on a manifold and on linear spaces is how to update the iterates for a given search direction. If the manifold is linear, its tangent space coincides with the manifold itself and the current iterate can be added to the search direction to obtain the update. If the manifold is nonlinear, the additive update has to be replaced by a suitable generalization. A natural idea on Riemannian manifolds would be to compute an update via the exponential map, i.e., via geodesics, but in many cases such exponential can be expensive to compute, therefore the use of cheaper surrogates, so called *retractions* is advocated in [AMS09].

These retractions have to satisfy certain consistency conditions and the weaker these conditions are, the more flexibly the retractions can be chosen. Based on these ideas, many algorithms of unconstrained optimization have been carried over to Riemannian manifolds, and have been analised in this framework [HT04, Lue72]. In general, the use of nonlinear retractions enables to exploit given nonlinear problem structure within an optimization algorithm. While this is true in particular for nonlinear manifolds, it may also sometimes be beneficial to use nonlinear retractions even in the case of linear spaces.

In coupled problems, mixed formulations, or optimal control of the above listed physical models, additional equality constraints occur, and thus one is naturally led to equality constrained optimization on manifolds. However, up to now optimization algorithms on manifolds have mainly been constructed for the unconstrained case. In constrast, not much research has been conducted on the construction of algorithms for equality constrained optimization on manifolds. A work in the field of shape optimization considers equality constraints on vector bundles [SSW15].

The subject of this work is the construction of an algorithm for equality constrained optimization on manifolds. In the problem setting we consider the manifolds X and Y and the problem

$$\min_{x \in X} f(x) \;\; s.t. \;\; c(x) = y. \tag{1}$$

Here $f : X \longrightarrow \mathbb{R}$ is a twice differentiable functional with suitable smoothness properties. The twice differentiable operator $c : X \longrightarrow Y$ maps from the manifold X to the manifold Y, and is a submersion.

In this work, particular focus is put on ways to exploit problem structure, and on invariance properties of the algorithm, extending the ideas of affine invariant Newton methods [Deu11]. Our point of departure is an *affine covariant composite step method* [LSW17] which was used to solve optimal control problems, involving finite strain elasticity [LSW14]. Composite steps are a very popular class of optimization methods for equality constrained problems, as can be seen in [CGT00] and the references therein. The algorithmic idea is to partition the optimization step δx into a normal step δn that improves feasibility and a tangential step that improves optimality:

$$\delta x = \delta n + \delta t : \qquad \delta t \in \ker c'(x), \qquad \delta n \in (\ker c'(x))^{\perp}$$

Close to a solution, δn and δt add up to a Lagrange-Newton step, and fast local convergence is obtained. Far away, the two substeps are suitably scaled to achieve global convergence. The method in [LSW17] is such a composite step method. Its main feature is the invariance under affine transformations of the codomain space of c, known as affine covariance. The invariance properties are also important for algorithms on manifolds, since they render them in a natural way, at least approximately, invariant under the choice of local coordinates.

We generalize the composite step method the case on manifolds in the following way. At a current iterate x_k we pullback both the objective f and the constraint mapping c to linear spaces through suitable retraction mappings obtaining maps, $\mathbf{f}$ and $\mathbf{c}$ with linear spaces T_xM and $T_{c(x)}N$ as domain and codomain, namely:

$$\mathbf{f} : T_xX \longrightarrow \mathbb{R} \qquad \mathbf{c} : T_xX \longrightarrow T_{c(x)}Y$$

this is followed by the computation of the normal $\delta n \in \ker \mathbf{c}'^{\perp}$ and tangential $\delta t \in \ker \mathbf{c}'$ steps, corrections that belong to linear spaces. A third correction $\delta s \in \ker \mathbf{c}'^{\perp}$ is computed and will serve as a way to avoid the Marathos effect. Once all corrections are computed, we update by using a retraction on the manifold X via:

$$x_+ = R_x^X(\delta t + \delta n + \delta s).$$

We study the influence of the retractions on the convergence of the algorithm. While the case of second order consistent retractions is relatively straightforward to analyse, the analysis of first order consistent retractions is more subtle, but still yields, after some algorithmic adjustments, local superlinear convergence of our algorithm. We put special emphasis on establishing rather weak assumptions on the smoothness of the retractions. We only assume a kind of second order directional differentiability property at the origin. This has important practical aspects, giving as much freedom for the implementation of the retractions as possible.

1.1 An affine invariant composite step method

In [LSW17] a composite step method for the solution of equality constrained optimization with partial differential equations has been proposed. We will briefly recapitulate its most important features. For details we refer to [LSW17]. There, in the problem setting, a Hilbert space $(X, \langle\cdot,\cdot\rangle)$ together with a reflexive Banach space P are considered in order to solve the following optimization problem

$$\min_{x\in X} f(x) \;\; s.t \;\; c(x) = 0. \tag{2}$$

The functional $f : X \longrightarrow \mathbb{R}$ is twice continuously Fréchet differentiable and the nonlinear operator $c : X \longrightarrow P^*$ maps into the dual space of P so it can model a differential equation in weak form:

$$c(x) = 0 \text{ in } P^* \iff c(x)v = 0 \text{ for all } v \in P. \tag{3}$$

The Lagrangian function L is given by

$$L(x, p) := f(x) + pc(x) \tag{4}$$

where the element p is the Lagrange multiplier at x. By $pc(x)$ we denote the dual pairing $P \times P^* \to \mathbb{R}$ with $pc(x) \in \mathbb{R}$. First and second derivatives of the Lagrangian function are:

$$L'(x, p) = f'(x) + pc'(x) \tag{5}$$

and

$$L''(x, p) = f''(x) + pc''(x). \tag{6}$$

In the *composite step method*, feasibility and optimality are carried out by splitting the full Lagrange-Newton step δx into a *normal step* δn and a *tangential step* δt. The normal step δn is a minimal norm Gauss-Newton step for the solution of the underdetermined problem $c(x) = 0$, and δt aims to minimize f on the current nullspace of the linearized constraints. For this, a *cubic regularization method* is employed. The following local problems are solved

$$\begin{aligned} \min_{\delta x} f(x) + f'(x)\delta x + \frac{1}{2}L''(x, p)(\delta x, \delta x) + \frac{[\omega_f]}{6}\|\delta x\|^3 \\ s.t.\ \nu c(x) + c'(x)\delta x = 0, \\ \frac{[\omega_c]}{2}\|\delta x\| \le \Theta_{aim}, \end{aligned}$$

where $\nu \in (0, 1]$ is an adaptively computed damping factor, $[\omega_{\mathbf{c}_2}]$ and $[\omega_{\mathbf{f}_2}]$ are algorithmic parameters, and Θ_{aim} is a user provided desired contraction factor. The parameters $[\omega_{\mathbf{c}_2}]$ and $[\omega_{\mathbf{f}_2}]$ are used for globalization of this optimization algorithm. They are used to quantify the mismatch between the quadratic model to be minimized and the nonlinear problem to be solved.

1.2 Computation of composite steps

Here we show how to compute the normal steps Δn, the Lagrange multiplier p_x and the tangential step δt, for the equality constrained problem in the linear setting. All these quantities are computed as solutions of certain saddle point problems. As a review, we present the way these quantities are computed, which also serves as a motivation for the manifold case, for more details see [LSW17].

In this section we suppose that $f : X \longrightarrow \mathbb{R}$ is twice continuously differentiable, X is a Hilbert space, $c(x) : X \longrightarrow P^*$ is a bounded, surjective twice differentiable mapping, and P is a reflexive space.

Normal step. It is well known that the minimal norm problem

$$\min_{v\in X} \frac{1}{2}\langle v, v\rangle \quad s.t \quad c'(x)v + g = 0, \tag{7}$$

is equivalent to the linear system

$$\begin{pmatrix} M & c'(x)^* \\ c'(x) & 0 \end{pmatrix} \begin{pmatrix} v \\ q \end{pmatrix} + \begin{pmatrix} 0 \\ g \end{pmatrix} = 0 \tag{8}$$

for some $g \in P^*$. Then, as shown in [LSW17], $v \in \ker c'(x)^\perp$. If the solution of the latter system is denoted as $v = -c'(x)^- g$, then we define the full normal step via

$$\Delta n := -c'(x)^- c(x).$$

For globalization, a damping factor $\nu \in]0, 1]$ is applied, setting $\delta n := \nu \Delta n$.

Lagrangian multiplier. At a point $x \in X$ we first compute a Lagrange multiplier p_x as the solution to the system:

$$\begin{pmatrix} M & c'(x)^* \\ c'(x) & 0 \end{pmatrix} \begin{pmatrix} v \\ p_x \end{pmatrix} + \begin{pmatrix} f'(x) \\ 0 \end{pmatrix} = 0. \tag{9}$$

It has been shown in [LSW17] that p_x is given uniquely, if $c'(x)$ is surjective, and p_x satisfies

$$f'(x)w + p_x c'(x)v = 0 \ \ \forall v \ \in \ \ker c'(x)^\perp.$$

This p_x will be called the Lagrange multiplier of the problem (2) at x.

Tangential step. With the help of p_x we define the quadratic model

$$q(\delta x) := f(x) + f'(x)\delta x + \frac{1}{2} L''(x, p_x)(\delta x, \delta x) \tag{10}$$

on $\ker c'(x)$. We solve the following quadratic problem in order to find the tangential step δt

$$\min_{\Delta t} q(\delta n + \Delta t) \quad \text{s. t.} \quad c'(x)\Delta t = 0. \tag{11}$$

which is equivalent to

$$\min_{\Delta t} \left(L'(x, p_x) + L''(x, p_x)\delta n\right) \Delta t + \frac{1}{2} L''(x, p_x)(\Delta t, \Delta t) \quad \text{s.t.} \quad c'(x)\Delta t = 0, \tag{12}$$

with corresponding first order optimality conditions

$$\begin{pmatrix} L''(x, p_x) & c'(x)^* \\ c'(x) & 0 \end{pmatrix} \begin{pmatrix} \Delta t \\ q \end{pmatrix} + \begin{pmatrix} L'(x, p_x) + L''(x, p_x)\delta n \\ 0 \end{pmatrix} = 0. \tag{13}$$

as long as L'' is positive definite on $\ker c'(x)$, which assures the existence of an exact minimizer. For the purpose of globalization, a cubic term is added to q, ensuring also existence of a minimizer, if positive definiteness fails. More details can be found in [LSW17].

Simplified normal step. For purpose of globalization and to avoid the Maratos effect, we compute a simplified normal step, which also plays the role of a second order correction.

The simplified Newton step is defined as

$$\delta s := -c'(x)^-(c(x + \delta x) - c(x) - c'(x)\delta x), \tag{14}$$

which ammount in solving a system of type (8). It can be seen from (8) that $\delta s \in \ker c'(x)^\perp$, and thus $(f'(x) + p_x c'(x))\delta s = 0$. It has been shown in [LSW17] that $f(x + \delta x + \delta s) - q(\delta x) = o(\|\delta x\|^2)$ is asymptotically more accurate than $f(x + \delta x) - q(\delta x) = O(\|\delta x\|^2)$. We will extend this result to the case of manifolds.

Update of iterates. If δx satisfies some acceptance criteria (cf. [LSW17]), the next iterate is computed as:

$$x_+ = x + \delta x + \delta s.$$

Of course, computation is only possible, because X is a linear space. To generalize our algorithm to manifolds, we have to replace this update by something different.

2 SQP-methods on a manifold

We generalize the *composite step method* from the setting of linear spaces, to the one in which the involved spaces are manifolds. Now we consider the problem

$$\min_{x \in X} f(x) \quad \text{s.t} \quad c(x) = y_*. \tag{15}$$

where the twice differentiable functional $f : X \longrightarrow \mathbb{R}$ is defined over the manifold X and the twice differentiable submersion $c : X \longrightarrow Y$ maps from the manifold X to the manifold Y. Further, $y_* \in Y$ is the required point.

Classical SQP-methods on vector spaces introduce local quadratic models for f and c at a given iterate x. In addition an SQP-method on a manifold has to provide local linear models for the nonlinear manifolds X and Y at x. From a differential geometric point of view, the tangent spaces T_xX and T_yY can be used for this purpose. Now local linear models for f and c can be defined as $T_xf : T_xX \to \mathbb{R}$ and $T_xc : T_xX \to T_{c(x)}Y$. However, quadratic approximations cannot be defined canonically. In differential geometry there are several ways to introduce additional structure to solve this problem. One well known example among these structures is a Riemannian metric, which allows the definition of geodesics and of the exponential map:

$$\exp_x : T_xX \to X$$

that locally maps each tangent vector $v \in T_xX$ to a geodesic, starting in x in direction v. Now pullbacks of f and c can be computed, and their corresponding first and second derivatives can be used to define quadratic models of f and c on T_xX and T_yY.

In this way, a quadratic optimization problem with linear constraints can be defined on T_xX and corresponding corrections δn, δt and p_x can be computed in a similar way as in Section 1.2 and also a trial step δx. By the exponential map a new iterate can be found via $x_+ = \exp_x(\delta x)$.

2.1 Consistency of retractions

However, often $\exp_x$ is hard or very expensive to evaluate, so in the optimization literature [AMS09], the notion of *retractions* has become customary, which can be seen as an efficient surrogate for $\exp_x$.

Definition 2.1. *A (first order) C^k-retraction ($k \geq 1$) on a manifold M is a mapping R^M from the tangent bundle TM onto M with the following properties. Let R^M_x denote the restriction of R^M to T_xM.*

i) $R^M_x(0_x) = x$, where 0_x denotes the zero element of T_xM.

ii) R^M_x is k-times continuously differentiable.

iii) With the canonical identification $T_{0_x}T_xM \simeq T_xM$, R^M_x satisfies

$$DR^M_x(0_x) = id_{T_xM}, \tag{16}$$

where id_{T_xM} denotes the identity mapping on T_xM.

If in addition $k \geq 2$ and

$$D^2R^M_x(0_x) = 0, \tag{17}$$

then R^M is called a retraction of second order.

More generally, it would be sufficient and appropriate to define a retraction only on a neighbourhood $U \subset T_xM$ of 0_x and not on all of T_xM. However, this would add additional technicalities to our study. For practical implementation in an optimization algorithm retractions should have a sufficiently large domain of definition, so that $R_x^M(\delta x)$ is defined for reasonable trial corrections δx. If necessary, $\delta x \in U$ can be enforced by additional scaling.

By the inverse mapping theorem R_x^M is locally continuously invertible and:

$$D(R_x^M)^{-1}(x) = (DR_x^M(0_x))^{-1} = id_{T_xM}.$$

In the following we consider a slightly different smoothness assumption on our retractions that is motivated from practical considerations.

Definition 2.2. *A first order C^1-retraction R^M is called a $C^{2,dir}$-retraction (second order directionally differentiable) if for each $v \in T_xM$ the mapping $t \to DR_x^M(tv) \in L(T_xM, T_xM)$ is differentiable with respect to t. We denote by $D^2R_x^M(v,w)$ the directional derivative of DR_x^M into direction v, applied to w.*

We observe that $D^2R_x^M(0_x)(v,w)$ is homogenous in v and w, and linear in w but not necessarily linear in v. This slightly weakened assumption, compared to C^2-retractions enables additional freedom in the choice and implementation of R^M. It is, for example possible to select different retractions, depending on the direction v as long as all of them are first order retractions. A very simple example for a $C^{2,dir}$-retraction on $M = \mathbb{R}$ would be

$$R_x^M(\delta x) := x + \delta x + \frac{\alpha}{2}\max\{\delta x, 0\}^2, \quad DR_0^M(0)v = v, \quad D^2R_0^M(0)(v,w) = \begin{cases} \alpha vw & : \quad v \geq 0 \\ 0 & : \quad v \leq 0 \end{cases}.$$

Certainly, the exponential map $R_x^M = \exp_x$ is the most prominent retraction of second order. Retractions can be considered as local approximations of the exponential map at a given point. Often, first order retractions are easier to compute than second order retractions. It is thus of interest, in how far algorithmic quantities depend on the choice of retraction. In the context of unconstrained optimization it is known (cf. e.g. [HT04, AMS09]) that first order retractions are sufficient in many aspects.

From a more general point of view, the construction of an SQP method involves a pair of retractions. One of them (e.g. the exponential map) is used to establish a quadratic model of the problem on the tangent space. The other retraction is used to compute the update $x_+ = R_x^M(\delta x)$. These two retractions can be *consistent* of first or second order. This frees us from the requirement to establish a Riemannian metric or compute covariant derivatives.

Definition 2.3. *On a smooth manifold M consider a pair of C^k-retractions at $x \in M$*

$$R_{x,i}^M : T_xM \to M \qquad i = 1, 2$$

and their local transformation mapping:

$$\Phi_M := (R_{x,1}^M)^{-1} \circ R_{x,2}^M : T_xM \to T_xM.$$

The pair $(R_{x,1}^M, R_{x,2}^M)$ of C^k-retractions is called first order consistent, if $k \geq 1$ and $\Phi'_M(0_x) = id_{T_xM}$ and second order consistent, if in addition $k \geq 2$ and $\Phi''_M(0_x) = 0$.

As a special case, a retraction R_x^M is of first (second) order in the sense of Definition 2.1, if it is consistent of first (second) order with $\exp_x$.

The following results for first order consistent C^1-retractions are easy to compute

$$\Phi_M(0_x) = 0_x \qquad \Phi'_M(0_x) = id_{T_xM},$$

For C^2-retractions we have in addition:

$$(\Phi_M^{-1})''(0_x) = -\Phi''_M(0_x).$$

The last result follows from the computation:

$$(\Phi_M^{-1})''(0_x) = [(\Phi'_M)^{-1}]'(0_x) = -(\Phi'_M)^{-1}(0_x)\Phi''_M(0_x)(\Phi'_M)^{-1}(0_x) = -\Phi''_M(0_x).$$

As a consequence we have the following results:

Lemma 2.1.

i) Every pair of first (second) order retractions is first (second) order consistent.

ii) $(R^M_{x,1}, R^M_{x,2})$ *is first (second) order consistent iff* $(R^M_{x,2}, R^M_{x,1})$ *is.*

The following case will play an important role in our work: if R^M_1 is a C^2 retraction and R^M_2 is a $C^{2,dir}$-retraction, then the mapping $(v, w) \to \Phi''_M(0_x)(v, w)$ is again linear in w and homogenous in v and w, but not necessarily linear in v.

These considerations lay the ground for the following section. First, we describe how to derive local quadratic models with the help of retractions and how to compute the substeps δn and δt on T_xM. Then we introduce the notion of consistency of a pair of retractions and discuss the consequences of this notion for SQP-algorithms. In particular, we will derive a quadratic model that is useful for a first order consistent pair of retractions.

Remark 2.1. *From a practical point of view, optimization algorithms on manifolds need not necessarily be based on the notion of tangent spaces and retractions. It is sufficient to define a local chart at each iterate, compute a local update in the chart with the help of a suitable quadratic model, and then perform the update by applying the local chart to the update. We will see in Section 5.4 below, that an implementation via local charts of* M *can be rather straightforward and convenient. From a conceptual point of view, however, working with tangent spaces and retractions is advantageous.*

2.2 The Lagrange function of the pulled-back problem

Next we will extend our SQP-algorithm to the case of manifolds, using retractions. For a given iterate $x \in X$ with $y = c(x) \in Y$ we have to perform two tasks:

1. Construct a linear-quadratic model of f and c on T_xX and T_yY. This will be done, using C^2-retractions $R^X_{x,1}$ and $R^Y_{y,1}$, as for example the exponential maps. These retractions need not be implemented, but serve as a way to derive linear and quadratic terms that make up the model. With the help to this model, a trial direction δx can be computed just as in the vector space case.

2. Given $\delta x \in T_xX$ compute an update that generalizes $x + \delta x$. This will be done, using a $C^{2,dir}$-retraction $R^X_{x,2}$ to obtain a new iterate $x_+ = R^X_{x,2}(\delta x)$. In addition, we need to evaluate in T_xY the preimage of $c(x_+)$ in T_yY with respect to a C^2-retraction $R^Y_{y,2}$. For that we need its inverse $(R^Y_{y,2})^{-1}$. Only $R^X_{x,2}$ and $(R^Y_{y,2})^{-1}$ have to be implemented.

The following assumptions will be taken:

Assumption 2.1. *Consider for* $x \in X$ *and* $y \in Y$ *the follwing first order consistent pairs of retractions:*

$$R^X_{x,i} : T_xX \to X \qquad i = 1, 2$$

and

$$R^Y_{y,i} : T_yY \to Y \qquad i = 1, 2,$$

where R^X_1, R^Y_1, *and* R^Y_2 *are* C^2*-retractions, and* R^X_2 *is a* $C^{2,dir}$*-retraction.*

Their local transformation mappings read:

$$\Phi_X := (R^X_{x,1})^{-1} \circ R^X_{x,2} : T_xX \to T_xX$$
$$\Phi_Y := (R^Y_{y,1})^{-1} \circ R^Y_{y,2} : T_yY \to T_yY.$$

We define the pull-back of the cost functional via the retraction:

$$\mathbf{f}_i : T_xX \longrightarrow \mathbb{R}$$
$$\mathbf{f}_i(u) = (f \circ R^X_{x,i})(u)$$

Similarly, we may pull-back the equality constraint operator $c : X \to Y$ locally:

$$c \circ R^X_{x,i} : T_x X \to Y.$$

To obtain a mapping $\mathbf{c}_i : T_x X \to T_y Y$ we have to define a push-forward via $R^Y_{y,i}$ as follows

$$\begin{aligned} \mathbf{c}_i : T_x X &\longrightarrow T_y Y \\ \mathbf{c}_i(u) &:= (R^Y_{y,i})^{-1} \circ c \circ R^X_{x,i}(u). \end{aligned}$$

The pullbacked mappings $\mathbf{f}_i$ and $\mathbf{c}_i$ are maps with linear spaces as domain and co-domain, therefore we are allowed to take first and second order derivatives in the usual way. This will be used throughout this work. We note, however, that these derivatives are only defined locally and may depend on the choice of retraction.

We can now define a local Lagrangian function via the pull-backs of f and c:

Definition 2.4. *The Lagrangian function at the point x with retractions R^X_x and R^Y_y is given by:*

$$\begin{aligned} \boldsymbol{L}_i(u,p) &= \boldsymbol{f}_i(u) + p\mathbf{c}_i(u) \\ &= f \circ R^X_{x,i}(u) + p(R^Y_{y,i})^{-1} \circ c \circ R^X_{x,i}(u) \end{aligned} \tag{18}$$

for $u \in T_x X$ and $p \in (T_y Y)^*$.

Observe that the dual pairing $p\mathbf{c}_i(u)$ is only possible, since $\mathbf{c}_i(u) \in T_y Y$ is the pull-back. A gobal definition of a Lagrangian function would require a nonlinear Lagrange multiplier $\tilde{p} : Y \to \mathbb{R}$.

For our purpose, we need to compute first and second derivatives of the Lagrangian function:

$$\mathbf{L}'_i(0_x, p_x)v := \mathbf{f}'_i(0_x)v + p_x \mathbf{c}'_i(0_x)v \tag{19}$$

$$\mathbf{L}''_i(0_x, p_x)(v,v) := \mathbf{f}''_i(0_x)(v,v) + p_x \mathbf{c}''_i(0_x)(v,v). \tag{20}$$

We observe that our definition of $\mathbf{L}$ is again a local one that depends on the given pair of retractions. In particular, we have:

$$\begin{aligned} \mathbf{L}_2(u,p) &= \mathbf{f}_2(u) + p\mathbf{c}_2(u) = \mathbf{f}_1 \circ \Phi_X(u) + p\Phi_Y^{-1} \circ \mathbf{c}_1 \circ \Phi_X(u) \\ &= \mathbf{L}_1 \circ \Phi_X(u) + p(\Phi_Y^{-1} - id) \circ \mathbf{c}_1 \circ \Phi_X(u). \end{aligned} \tag{21}$$

Differentiating this expression at 0_x, using the chain rule, we obtain the identities:

$$\mathbf{f}'_1(0_x) = \mathbf{f}'_2(0_x), \qquad \mathbf{c}'_1(0_x) = \mathbf{c}'_2(0_x), \qquad \mathbf{L}'_1(0_x, p) = \mathbf{L}'_2(0_x, p).$$

Hence, we do not need to distinguish and thus we use the notation $\mathbf{f}'(0_x)$, $\mathbf{c}'(0_x)$, $\mathbf{L}'(0_x, p)$. However, concerning $\mathbf{L}''_i$ we obtain different expressions. In particular, while $\mathbf{L}''_1$ is a bilinear form, $\mathbf{L}''_2$ may be not, because $R^X_{2,x}$ is only a $C^{2,dir}$ retraction.

Lemma 2.2.

$$(\mathbf{L}''_2(0_x, p_x) - \mathbf{L}''_1(0_x, p_x))(v,w) = \mathbf{L}'(0_x, p_x)\Phi''_X(0_x)(v,w) - p_x \Phi''_Y(0_y)(\mathbf{c}'(0_x)v, \mathbf{c}'(0_x)w). \tag{22}$$

In particular:

i) if $(R^X_{x,1}, R^X_{x,2})$ is second order consistent, or $\mathbf{L}'(0_x, p_x) = 0$, then $\mathbf{L}''_1(0_x, p_x) = \mathbf{L}''_2(0_x, p_x)$ on $\ker \mathbf{c}'(0_x)$.

ii) if $(R^X_{x,1}, R^X_{x,2})$ and $(R^Y_{y,1}, R^Y_{y,2})$ are second order consistent, then $\mathbf{L}''_1(0_x, p_x) = \mathbf{L}''_2(0_x, p_x)$ on $T_x X$.

Proof. We compute by the chain rule:

$$\begin{aligned}\mathbf{f}_2''(0_x)(v,w) - \mathbf{f}_1''(0_x)(v,w) &= \mathbf{f}'(0_x)\Phi_X''(0_x)(v,w)\\ \mathbf{c}_2''(0_x)(v,w) - \mathbf{c}_1''(0_x)(v,w) &= (\Phi_Y^{-1})''(0_y)(\mathbf{c}'(0_x)v, \mathbf{c}'(0_x)w) + \mathbf{c}'(0_x)\Phi_X''(0_x)(v,w)\\ &= -\Phi_Y''(0_y)(\mathbf{c}'(0_x)v, \mathbf{c}'(0_x)w) + \mathbf{c}'(0_x)\Phi_X''(0_x)(v,w).\end{aligned} \tag{23}$$

□

Remark 2.2. *Obviously,* $\mathbf{L}_1''(0_x,p_x)(v,w) = \mathbf{L}_2''(0_x,p_x)(v,w)$ *if* x *is a KKT-point, i.e.,* $\mathbf{L}'(0_x,p_x)=0$ *and* v *or* $w \in \ker \mathbf{c}'(0_x)$*. Hence, second order optimality conditions are invariant under change of retractions. This is, of course, to be expected.*

Moreover, close to a KKT point, $\mathbf{L}_1''(0_x,p_x) - \mathbf{L}_2''(0_x,p_x)$ *is small on* $\ker \mathbf{c}'(0_x)$*. Thus, if* x *is an SSC point, we obtain invertibility of the Lagrange-Newton matrix in a neighbourhood of* x*, regardless of the choice of retraction.*

2.3 Computation of the steps

The computation of the normal and tangential corrections as well as the Lagrange multiplier are done in a similar way as in the linear case. First, the mappings are pullbacked to linear spaces through the local parametrizations and there, we compute the quantities as solution of certain saddle point problems.

Normal step. We note that the minimal norm problem

$$\min_{w\in T_xX} \frac{1}{2}\langle w,w\rangle \ s.t.\ \mathbf{c}'(0_x)w + g = 0, \tag{24}$$

is equivalent to finding $w \in \ker \mathbf{c}'(0_x)^\perp$ such that $\mathbf{c}'(0_x)w + g = 0$ and we write in short $w = -\mathbf{c}'(0_x)^- g$.

Let $\mathbf{M}_x : T_xX \to (T_xX)^*$ given via a scalar product $(\mathbf{M}_x v)w = \langle v,w\rangle_x$ (possibly depending on x) and thus symmetric and positive definite. If, for example, a Riemmannian metric is given on X, then $\langle v,w\rangle_x$ may be chosen as the corresponding scalar product. Then the system:

$$\begin{pmatrix} \mathbf{M}_x & \mathbf{c}'(0_x)^* \\ \mathbf{c}'(0_x) & 0 \end{pmatrix}\begin{pmatrix} w \\ q \end{pmatrix} + \begin{pmatrix} 0 \\ g \end{pmatrix} = 0 \tag{25}$$

corresponds to the KKT-conditions for (24), and thus the solutions of (25) and (24) concide.

Now we can define the full normal step as follows:

$$\Delta n := -\mathbf{c}'(0_x)^-(\mathbf{c}(0_x) - \mathbf{y}).$$

as solution of (25) and (24) with $g = \mathbf{c}(0_x) - \mathbf{y}_*$, where $\mathbf{y}_* = (R_y^Y)^{-1}(y_*)$. For globalization we will use damped normal steps $\delta n := \nu \Delta n$ with a damping factor $\nu \in]0,1]$.

Lagrangian multiplier. The Lagrange multiplier is the element p_x that solves

$$\begin{pmatrix} \mathbf{M}_x & \mathbf{c}'(0_x)^* \\ \mathbf{c}'(0_x) & 0 \end{pmatrix}\begin{pmatrix} w \\ p_x \end{pmatrix} + \begin{pmatrix} \mathbf{f}' \\ 0 \end{pmatrix} = 0$$

and the latter implies that p_x satisfies

$$\mathbf{f}'(0_x)v + p_x\mathbf{c}'(0_x)v = 0 \quad \forall v \in \ker \mathbf{c}'(0_x)^\perp. \tag{26}$$

Note that p_x is a linear function:

$$p_x : T_{\mathbf{c}(0_x)}Y \longrightarrow \mathbb{R}$$

i.e., $p_x \in T_{\mathbf{c}(0_x)}Y^*$. It can be observed easily that p_x is independent of the choice of first order retraction, as long as $\mathbf{M}_x$ does not change.

Tangential step. Up to now, the computed quantities do not depend on the choice of retraction. However, the tangent step will. After computing Δn a damping factor ν, such that $\delta n = \nu \Delta n$, and an adjoint state p_x, we compute the tangential step $\delta t \in \ker \mathbf{c}'(0_x)$.

Using (19) and (20) we define the quadratic model as:

$$\mathbf{q}_1(\delta x) := \mathbf{f}(0_x) + \mathbf{f}'(0_x)\delta x + \frac{1}{2}\mathbf{L}_1''(0_x, p_x)(\delta x, \delta x),$$

if $\delta x := \delta n + \Delta t$ with $\Delta t \in \ker \mathbf{c}'(0)$ and $\delta n \in \ker \mathbf{c}'(0)^\perp$ then

$$\mathbf{q}_1(\delta x) = \mathbf{f}(0_x) + \mathbf{f}'(0_x)(\Delta t + \delta n) + \frac{1}{2}\mathbf{L}_1''(0_x, p_x)(\Delta t + \delta n, \Delta t + \delta n)$$

For given $\delta n = \nu \Delta n$ the tangential step δt is found by solving approximately the problem

$$\min_{\Delta t} \mathbf{q}_1(\delta n + \Delta t) \;\; s.t \;\; \mathbf{c}'(0_x)\Delta t = 0,$$

which, after adding the term $p_x \mathbf{c}'(0_x)\Delta t = 0$ and omitting terms that are independent of δt is equivalent to:

$$\begin{aligned} &\min_{\Delta t} \left(\mathbf{L}'(0_x, p_x) + \mathbf{L}_1''(0_x, p_x)\delta n\right)\Delta t + \frac{1}{2}\mathbf{L}_1''(0_x, p_x)(\Delta t, \Delta t) \\ &\text{s.t.} \quad \mathbf{c}'(0_x)\Delta t = 0. \end{aligned}$$

By assumption, since R_1^X is a C^2-retraction this is a quadratic problem that can be solved by standard means. Of course, in the presence of non-convexity an exact solution does not always exist, but there are various algorithmic ways (e.g. truncated cg) to compute an appropriate surrogate. In contrast, using only a $C^{2,dir}$-retraction would lead to a nonlinear minimization problem at this point, which would be much harder to solve.

Close to a solution satisfying the second order conditions ($\mathbf{L}''$ positive definite on $\ker \mathbf{c}'$) then the solution to the previous problem exists, and the first order optimality conditions are

$$\begin{pmatrix} \mathbf{L}_1''(0_x, p_x) & \mathbf{c}'(0_x)^* \\ \mathbf{c}'(0_x) & 0 \end{pmatrix} \begin{pmatrix} \Delta t \\ q \end{pmatrix} + \begin{pmatrix} \mathbf{L}'(0_x, p_x) + \mathbf{L}_1''(0_x, p_x)\delta n \\ 0 \end{pmatrix} = 0. \tag{27}$$

Again, for purpose of globalization we may compute a different tangent step δt (using, for example a line-search, a trust-regions, or cubic regularization), and set $\delta x = \delta n + \delta t$.

Simplified normal step. In the same way as above, a simplified normal step can be computed via

$$\delta s := -\mathbf{c}'(0_x)^- \left(\mathbf{c}_2(\delta x) - \mathbf{c}(0_x) - \mathbf{c}'(0_x)\delta x\right),$$

which is used for our globalization mechanism and as a second order correction. For the computation of δs, we have to evaluate

$$\mathbf{c}_2(\delta x) = (R_{y,2}^Y)^{-1} \circ c \circ R_{x,2}^X(\delta x).$$

This is possible, because $R_{x,2}^X$ and $(R_{y,2}^Y)^{-1}$ are implemented. Since this is not the case for $R_{x,1}^X$ and $(R_{y,1}^Y)^{-1}$ it would not be possible to evaluate $\mathbf{c}_1(\delta x)$.

Updates of iterates. As already noted before, new iterates are computed using $R_{x,2}^X$, namely:

$$x_+ := R_{x,2}^X(\delta x + \delta s).$$

Thus, for the new objective function value, we obtain:

$$f(x_+) = f(R_{x,2}^X(\delta x + \delta s)) = \mathbf{f}_2(\delta x + \delta s).$$

2.4 Consistency of quadratic models

To study invariance, we consider the case that our local model, depending on $\mathbf{f}$, $\mathbf{c}$, and its first and second derivatives, is computed with respect to the C^2-retractions R_1^X and R_1^Y, while the actual evaluation of f and c are performed with respect to the $C^{2,dir}$ retraction R_2^X and the C^2-retraction R_2^Y. We assume only first order consistency of (R_1^X, R_2^X) and (R_1^Y, R_2^Y).

Lemma 2.3. *For a given perturbation $\delta x \in T_x X$ let $\delta s \in \ker \mathbf{c}'(0_x)^\perp$ be the simplified normal step, given by the minimal norm solution of the equation:*

$$-\boldsymbol{c}'(0_x)\delta s = \boldsymbol{c}_2(\delta x) - \boldsymbol{c}(0_x) - \boldsymbol{c}'(0_x)\delta x. \tag{28}$$

Then the following identity holds:

$$\boldsymbol{f}_2(\delta x + \delta s) - \boldsymbol{q}_1(\delta x) = \boldsymbol{r}_2(\delta x) + \boldsymbol{s}_2(\delta x) + \frac{1}{2}\left(\boldsymbol{L}'(0_x, p_x)\Phi_X''(\delta x, \delta x) - p_x \Phi_Y''(\boldsymbol{c}'(0_x)\delta n, \boldsymbol{c}'(0_x)\delta n)\right). \tag{29}$$

where

$$\begin{aligned}
\boldsymbol{r}_2(\delta x) &:= \boldsymbol{L}_2(\delta x, p_x) - \boldsymbol{L}(0_x, p_x) - \boldsymbol{L}'(0_x, p_x)\delta x - \frac{1}{2}\boldsymbol{L}_2''(0_x, p_x)(\delta x, \delta x) \\
\boldsymbol{s}_2(\delta x) &:= \boldsymbol{f}_2(\delta x + \delta s) - \boldsymbol{f}_2(\delta x) - \boldsymbol{f}'(0_x)\delta s.
\end{aligned}$$

in addition, we have:

$$\delta s = \int_0^1 \boldsymbol{c}'(0_x)^-(\boldsymbol{c}_2'(\sigma\delta x) - \boldsymbol{c}'(0_x))\delta x \, d\sigma. \tag{30}$$

Proof. Using the fundamental theorem of calculus, from (28) we get (30). In order to proof (29), we start with

$$\begin{aligned}
\mathbf{r}_2(\delta x) + \mathbf{q}_1(\delta x) &= \mathbf{L}_2(\delta x, p_x) - \mathbf{L}(0_x, p_x) - \mathbf{L}'(0_x, p_x)\delta x - \frac{1}{2}\mathbf{L}_2''(0_x, p_x)(\delta x, \delta x) \\
&\quad + \mathbf{f}(0_x) + \mathbf{f}'(0_x, p_x)\delta x + \frac{1}{2}\mathbf{L}_1''(0_x, p_x)(\delta x, \delta x) \\
&= \mathbf{f}_2(\delta x) + p_x[\mathbf{c}_2(\delta x) - \mathbf{c}(0_x) - \mathbf{c}'(0_x)\delta x] + \frac{1}{2}\left(\mathbf{L}_1''(0_x, p_x) - \mathbf{L}_2''(0_x, p_x)\right)(\delta x, \delta x) \\
&= \mathbf{f}_2(\delta x) - p_x\mathbf{c}'(0_x)\delta s - \frac{1}{2}\left(\mathbf{L}'(0_x, p_x)\Phi_X''(\delta x, \delta x) - p_x\Phi_Y''(\mathbf{c}'(0_x)\delta x, \mathbf{c}'(0_x)\delta x)\right),
\end{aligned}$$

where the identity (22) has been used. Given that $\mathbf{f}'(0_x)\delta s = -p_x\mathbf{c}'(0_x)\delta s$ and adding and subtracting $\mathbf{f}_2(\delta x + \delta s)$, we obtain

$$\begin{aligned}
\mathbf{r}_2(\delta x) + \mathbf{q}_1(\delta x) &= \mathbf{f}_2(\delta x + \delta s) - \mathbf{f}_2(\delta x + \delta s) + \mathbf{f}_2(\delta x) + \mathbf{f}'(0_x)\delta s \\
&\quad - \frac{1}{2}\left(\mathbf{L}'(0_x, p_x)\Phi_X''(\delta x, \delta x) - p_x\frac{1}{2}\Phi_Y''(\mathbf{c}'(0_x)\delta x, \mathbf{c}'(0_x)\delta x\right).
\end{aligned}$$

Using finally $\mathbf{c}'(0_x)\delta x = \mathbf{c}'(0_x)\delta n$ we obtain (29). □

We observe that the difference of $\mathbf{f}_2$ to $\mathbf{q}_1$ is now second order, and not, as desired, of third order. There are two terms involved:

- The first term $\mathbf{L}'(0_x, p_x)\Phi_X''(\delta x, \delta x)$ is due to lack of second order consistency of Φ_X. We observe that this term vanishes at a KKT point and is small in a neighbourhood thereof.
- The second term $p_x\Phi_Y''(\mathbf{c}'(0_x)\delta x, \mathbf{c}'(0_x)\delta x)$ only affects normal directions, but it does not vanish at a KKT point. So it may affect the acceptance criteria of a globalization scheme and slow down local convergence.

2.5 A second order quadratic model for first order retractions

In the following we consider again first order consistent pairs of retractions. Taking into account that Φ_Y does not influence the computation of the steps, but may have negative effects on the globalization scheme, we look for an alternative to the quadratic model $\mathbf{q}_1$ with better consistency properties. Here we have to keep in mind that $\mathbf{L}_2''(0_x, p_x)$ is not available.

If (R_1^Y, R_2^Y) is second order consistent, then we use $\mathbf{q}_1$ as a model. However, the case when (R_1^Y, R_2^Y) is only first order consistent, we propose to give the following surrogate model:

$$\begin{aligned}\tilde{\mathbf{q}}(\delta n)(\delta t) &:= \mathbf{L}_2(\delta n, p_x) - (1-\nu)p_x\mathbf{c}(0_x) + (\mathbf{f}'(0_x) + \mathbf{L}_1''(0_x, p)\delta n)\delta t + \frac{1}{2}\mathbf{L}_1''(0_x, p)(\delta t, \delta t)\\ &= \mathbf{f}_2(\delta n) + p_x(\mathbf{c}_2(\delta n) - (1-\nu)\mathbf{c}(0_x)) + (\mathbf{f}'(0_x) + \mathbf{L}_1''(0_x, p)\delta n)\delta t + \frac{1}{2}\mathbf{L}_1''(0_x, p)(\delta t, \delta t).\end{aligned} \tag{31}$$

With this, we will show below:

$$\mathbf{f}_2(\delta x + \delta s) - \tilde{\mathbf{q}}(\delta n)(\delta t) = \frac{1}{2}\mathbf{L}'(0_x, p_x)(\Phi_X''(\delta x, \delta x) - \Phi_X''(\delta n, \delta n)) + o(\|\delta x\|^2).$$

Close to a KKT-point, the remaining second order term is small. It turns out that such a model is sufficient to show local superlinear convergence. The evaluation of $\tilde{\mathbf{q}}(\delta n)(\delta t)$ requires the evaluation of $\mathbf{L}_2(\delta n, p_x)$ which has to be done once per outer iteration. If $\nu < 1$, which is the case far away from a feasible point, $\mathbf{q}_1$ is used as a model.

Lemma 2.4. *For the surrogate model $\tilde{\mathbf{q}}$, we have that:*

$$\tilde{\mathbf{q}}(0_x, \delta n)(\delta t) - \boldsymbol{q}_1(\delta x) = r_2(\delta n) + \frac{1}{2}\left(\boldsymbol{L}'(0_x, p_x)\Phi_X''(\delta n, \delta n) - p_x\Phi_Y''(\boldsymbol{c}'(0_x)\delta n, \boldsymbol{c}'(0_x)\delta n)\right). \tag{32}$$

In particular, for fixed δn:

$$\operatorname*{argmin}_{\delta t \in \ker \mathbf{c}'(0_x)} \tilde{\mathbf{q}}(\delta n)(\delta t) = \operatorname*{argmin}_{\delta t \in \ker \mathbf{c}'(0_x)} \mathbf{q}_1(\delta n + \delta t).$$

Proof. By definition of $\mathbf{q}_1(v)$ we obtain, using the fact that $\nu p_x\mathbf{c}(0_x) = -p_x\mathbf{c}'(0_x)\delta n = \mathbf{f}'(0_x)\delta n$

$$\begin{aligned}&\mathbf{L}_2(\delta n, p_x) - \mathbf{L}(0_x, p_x) + \mathbf{q}_1(\delta x) - \frac{1}{2}\mathbf{L}_1''(0_x, p_x)\delta n^2\\ &\quad= \mathbf{L}_2(\delta n, p_x) - \mathbf{f}(0_x) - p_x\mathbf{c}(0_x) + \mathbf{f}(0_x) + \mathbf{f}'(0_x)\delta x + \frac{1}{2}\mathbf{L}_1''(0_x, p_x)\delta x^2 - \frac{1}{2}\mathbf{L}_1''(0_x, p_x)\delta n^2\\ &\quad= \mathbf{L}_2(\delta n, p_x) + (\nu - 1)p_x\mathbf{c}(0_x) + \mathbf{f}'(0_x)\delta t + \frac{1}{2}\mathbf{L}_1''(0_x, p_x)(\delta x + \delta n, \delta t) = \tilde{\mathbf{q}}(\delta n)(\delta t).\end{aligned}$$

Taking into account

$$\mathbf{L}_2(\delta n, p_x) - \mathbf{L}(0_x, p_x) = \mathbf{r}_2(\delta n) + \mathbf{L}'(0_x, p_x)\delta n + \frac{1}{2}\mathbf{L}_2''(0_x, p_x)\delta n^2 = \mathbf{r}_2(\delta n) + \frac{1}{2}\mathbf{L}_2''(0_x, p_x)\delta n^2$$

and (22) we obtain (32). □

Lemma 2.5. *For the surrogate model $\tilde{\mathbf{q}}$, we have the identity*

$$\boldsymbol{f}_2(\delta x + \delta s) - \tilde{\mathbf{q}}(\delta n)(\delta t) = \boldsymbol{r}_2(\delta x) - \boldsymbol{r}_2(\delta n) + \boldsymbol{s}_2(\delta x) + \frac{1}{2}\boldsymbol{L}'(0_x, p_x)(\Phi_X''(\delta x, \delta x) - \Phi_X''(\delta n, \delta n)). \tag{33}$$

Proof. By Lemma 2.3 and Lemma 2.4 we compute

$$\begin{aligned}\mathbf{f}_2(\delta x + \delta s) - \tilde{\mathbf{q}}(\delta n)(\delta t) &= (\mathbf{f}_2(\delta x + \delta s) - \mathbf{q}_1(\delta x)) - (\tilde{\mathbf{q}}(\delta n)(\delta t) - \mathbf{q}_1(\delta x))\\ &= \mathbf{r}_2(\delta x) + \mathbf{s}_2(\delta x) + \frac{1}{2}\left(\mathbf{L}'(0_x, p_x)\Phi_X''(\delta x, \delta x) - p_x\Phi_Y''(\mathbf{c}'(0_x)\delta n, \mathbf{c}'(0_x)\delta n)\right)\\ &\quad -\mathbf{r}_2(\delta n) - \frac{1}{2}\left(\mathbf{L}'(0_x, p_x)\Phi_X''(\delta n, \delta n) - p_x\Phi_Y''(\mathbf{c}'(0_x)\delta n, \mathbf{c}'(0_x)\delta n)\right)\\ &= \mathbf{r}_2(\delta x) + \mathbf{s}_2(\delta x) - \mathbf{r}_2(\delta n) + \frac{1}{2}\mathbf{L}'(0_x, p_x)\left(\Phi_X''(\delta x, \delta x) - \Phi_X''(\delta n, \delta n)\right).\end{aligned}$$

The crucial observation is that $p_x\Phi_Y''(\mathbf{c}'(0_x)\delta n, \mathbf{c}'(0_x)\delta n)$ cancels out. □

To quantify the remainder terms, we have to use quantitative assumptions on the nonlinearity of the problem and the retractions.

Proposition 2.1. *Assume that there are constants $\omega_{\boldsymbol{c}_2}$, $\omega_{\boldsymbol{f}_2'}$ and $\omega_{\boldsymbol{L}_2}$ such that*

$$\|\boldsymbol{c}'(0_x)^-(\boldsymbol{c}_2'(v)-\boldsymbol{c}'(0_x))w\| \le \omega_{\boldsymbol{c}_2}\|v\|\|w\|, \tag{34}$$

$$|(\boldsymbol{L}_2''(v,p_x)-\boldsymbol{L}_2''(0_x,p_x))(v,w)| \le \omega_{\boldsymbol{L}_2}\|v\|^2\|w\|, \tag{35}$$

$$|(\boldsymbol{f}_2'(v)-\boldsymbol{f}'(0_x))w| \le \omega_{\boldsymbol{f}_2'}\|v\|\|w\| \tag{36}$$

i.e. Lipschitz conditions holds for the pullback mappings with retraction R_2^X and R_2^Y, where v and w are arbitrary. Then for arbitrary δx and simplified normal step δs as defined in (28) *we have the estimates:*

$$\|\delta s\| \le \frac{\omega_{\boldsymbol{c}_2}}{2}\|\delta x\|^2$$

$$|\boldsymbol{f}_2(\delta x+\delta s)-\tilde{\mathbf{q}}(0_x,\delta n)(\delta t)| \le \Big(\frac{\omega_{\boldsymbol{L}_2}}{3}+\frac{\omega_{\boldsymbol{f}_2'}\omega_{\boldsymbol{c}_2}}{2}(1+\frac{\omega_{\boldsymbol{c}_2}}{4}\|\delta x\|)\Big)\|\delta x\|^3+\frac{1}{2}|\boldsymbol{L}'(0_x,p_x)(\Phi_X''(\delta x^2)-\Phi_X''(\delta n^2))|$$

Proof. By Assumption 2.1 all stated derivatives exist. In particular $\mathbf{L}_2''(v,p_x)(v,w)$ exists as a directional derivative of $\mathbf{L}_2'(v,p_x)w$ in direction v, since R_2^X is a $C^{2,dir}$-retraction. This is all we need in the following.

From (30), setting $v=\sigma\delta x$, we have that

$$\|\delta s\| \le \int_0^1 \frac{1}{\sigma}\|\mathbf{c}'(0_x)^-(\mathbf{c}_2'(\sigma\delta x)-\mathbf{c}'(0_x))\sigma\delta x\|\,d\sigma \le \frac{\omega_{\mathbf{c}_2}}{2}\|\delta x\|^2$$

by Lemma 2.5 we get

$$|\mathbf{f}_2(\delta x+\delta s)-\tilde{\mathbf{q}}(\delta n)(\delta t)| \le |\mathbf{r}_2(\delta x)|+|\mathbf{r}_2(\delta n)|+|\mathbf{s}_2(\delta x)|+\frac{1}{2}|\mathbf{L}'(0_x,p_x)(\Phi_X''(\delta x)^2-\Phi_X''(\delta n)^2)|.$$

Assuming the affine covariant Lipschitz conditions, we get that

$$|\mathbf{r}_2(v)| \le \int_0^1\int_0^1 \frac{1}{\tau^2\sigma}|(\mathbf{L}_2''(\tau\sigma v,p_x)-\mathbf{L}_2''(0_x,p_x))(\tau\sigma v,\tau\sigma v)|\,d\tau d\sigma \le \omega_{\mathbf{L}_2}\|v\|\int_0^1\int_0^1 \tau\sigma^2\,d\tau d\sigma = \frac{\omega_{\mathbf{L}_2}}{6}\|v\|^3$$

v is arbitrary, then the latter hold for $v=\delta x$ and $v=\delta n$

$$|\mathbf{r}_2(\delta x)|+|\mathbf{r}_2(\delta n)| \le \frac{\omega_{\mathbf{L}_2}}{6}\|\delta x\|^3+\frac{\omega_{\mathbf{L}_2}}{6}\|\delta n\|^3 \le \frac{\omega_{\mathbf{L}_2}}{3}\|\delta x\|^3$$

and for $\mathbf{s}_2$ we obtain

$$\begin{aligned}|\mathbf{s}_2(\delta x)| &\le \int_0^1 |(\mathbf{f}_2'(\delta x+\sigma\delta s)-\mathbf{f}'(0_x)\delta s)|d\sigma \le \omega_{\mathbf{f}_2'}\|\delta s\|\int_0^1\|\delta x+\sigma\delta s\|\,d\sigma\\ &\le \omega_{\mathbf{f}_2'}\|\delta s\|\left(\|\delta x\|+\frac{1}{2}\|\delta s\|\right) \le \frac{\omega_{\mathbf{f}_2'}\omega_{\mathbf{c}_2}}{2}\|\delta x\|^2\left(\|\delta x\|+\frac{\omega_{\mathbf{c}_2}}{4}\|\delta x\|^2\right)\end{aligned}$$

Adding all estimates up, we obtain the desired estimate. □

3 Globalization Scheme

In [LSW17, Section 4] a globalization scheme has been proposed for an affine covariant composite step method. In the following we will recapitulate its main features and adjust it to the case of manifolds, where necessary. Since our aim is to study local convergence of our algorithm, we concentrate on the aspects of our scheme that are relevant for local convergence.

Each step of the globalization scheme at a current iterate x will be performed on T_xX and T_yY, using $R_{x,i}^X$ and $R_{y,i}^Y$ as retractions to pull f and c back to T_xX and T_yY, as sketched in the previous section. Then the globalization scheme from [LSW17] can be used.

For given algorithmic parameters $[\omega_{\mathbf{f}_2}]$ and $[\omega_{\mathbf{c}_2}]$ and given damping-parameters ν, we compute the new trial correction δx as follows after Δn, p_x, Δt, ν have been computed.

$$\begin{aligned} \min_{\tau:\delta x=\nu\Delta n+\tau\Delta t} \mathbf{f}(0_x)+\mathbf{f}'(0_x)\delta x+\frac{1}{2}\mathbf{L}_1''(0_x,p_x)(\delta x,\delta x)+\frac{[\omega_{\mathbf{f}_2}]}{6}\|\delta x\|^3 \\ s.t.\ \nu\mathbf{c}(0_x)+\mathbf{c}'(0_x)\delta x=0, \\ \frac{[\omega_{\mathbf{c}_2}]}{2}\|\delta x\|\le\Theta_{\text{aim}}, \end{aligned} \tag{37}$$

With the restriction $\delta x=\nu\Delta n+\tau\Delta t$ this problem is actually a scalar problem in τ, which is simple to solve. More sophisticated strategies to compute δt directly as an approximate minimizer of the cubic model are conceivable and have been described in the literature.

Algorithm 1 Outer and inner loop (inner loop simplified)

Require: initial iterate x, $[\omega_{\mathbf{c}_2}], [\omega_{\mathbf{f}_2}]$
  repeat// *NLP loop*
    choose retractions $R^X_{x,2}$, $R^Y_{y,2}$ at x and y
    compute quadratic models of f and c, based on $R^X_{x,1}$ and $R^Y_{y,1}$
    repeat// *step computation loop*
      compute Δn, p_x
      compute maximal $\nu\in]0,1]$, such that $\frac{[\omega_{\mathbf{c}_2}]}{2}\|\nu\Delta n\|\le\rho_{\text{ellbow}}\Theta_{\text{aim}}$
      compute Δt via (27)
      compute trial correction δx, via (37)
      compute simplified correction δs, via (28)
      evaluate acceptance tests (38) and (40)
      compute new Lipschitz constants $[\omega_{\mathbf{c}_2}], [\omega_{\mathbf{f}_2}]$, using δs, $\mathbf{f}_2(\delta x+\delta s)$, and $\mathbf{q}_1(\delta x)$ or $\tilde{\mathbf{q}}(\delta n)(\delta t)$
    until trial correction δx accepted
    $x\leftarrow R^X_{x,2}(\delta x+\delta s)$
  until converged

As elaborated in [LSW17] we use the algorithmic parameter $[\omega_{\mathbf{c}_2}]$ to capture the nonlinearity of c, while $[\omega_{\mathbf{f}_2}]$ models the nonlinearity of f. Initial estimates have to be provided.

After computation of Δn, we compute a maximal damping factor $\nu\in]0,1]$ and $\delta n:=\nu\Delta n$, such that

$$\frac{[\omega_{\mathbf{c}_2}]}{2}\|\delta n\|\le\rho_{\text{ellbow}}\Theta_{\text{aim}}.$$

Here $\Theta_{\text{aim}}\in]0,1[$ is a desired Newton contraction for the underdetermined problem $\mathbf{c}_2(x)=0$ and $\rho_{\text{ellbow}}\in]0,1]$ provides some ellbow space in view of the last line of (37), which can be seen as a trust-region constraint, governed by the nonlinearity of c.

Then, Δt is computed via (27). If $\mathbf{L}_1''$ is not positive definite on $\ker\mathbf{c}'(0_x)$, then a suitable modified solution (e.g. form truncated cg) is used. Then

$$\delta x:=\delta n+\tau\Delta t$$

is computed via minimizing (37) over τ and the simplified normal step δs is computed via (28).

At this point updates for $[\omega_{\mathbf{c}_2}]$ and $[\omega_{\mathbf{f}_2}]$ can be computed. Just as in [LSW17] we define

$$[\omega_{\mathbf{c}_2}]:=\frac{2\|\delta s\|}{\|\delta x\|^2}$$

as an affine covariant quantity that measures the nonlinearity of c. Concerning $[\omega_{\mathbf{f}_2}]$, the use of retractions requires a modification, compared to [LSW17]. We first define

$$\mathbf{q}(\delta x):=\begin{cases}\mathbf{q}_1(\delta x) & : \ (R^Y_1,R^Y_2) \text{ is second order consistent}\\ \tilde{\mathbf{q}}(\delta n)(\delta t) & : \ \text{otherwise}\end{cases}$$

and then set:

$$[\omega_{\mathbf{f}_2}]^{\text{raw}} := \frac{6}{\|\delta x\|^3}(\mathbf{f}_2(\delta x + \delta s) - \mathbf{q}(\delta x)).$$

This (potentially negative) estimate has to be augmented by some save-guard bounds of the form

$$[\omega_{\mathbf{f}_2}]^{\text{new}} = \min\{\rho_1[\omega_{\mathbf{f}_2}]^{\text{old}}, \max\{\rho_0[\omega_{\mathbf{f}_2}]^{\text{old}}, [\omega_{\mathbf{f}_2}]^{\text{raw}}\}\}$$

with $0 < \rho_0 < 1 < \rho_1$.

For acceptance of iterates, we perform a contraction test and a decrease test. The contraction test requires, just as in [LSW17],

$$\frac{\|\delta s\|}{\|\delta x\|} \leq \Theta_{\text{acc}} \tag{38}$$

for acceptance, with some parameter $\Theta_{\text{acc}} \in]\Theta_{\text{des}}, 1[$. For the decrease test we define

$$\mathbf{m}_{[\omega_{\mathbf{f}_2}]}(v) := \mathbf{q}(v) + \frac{[\omega_{\mathbf{f}_2}]}{6}\|v\|^3.$$

and require some ratio of actual decrease and predicted decrease condition. We choose $\underline{\eta} \in]0, 1[$ and define:

$$\eta := \frac{\mathbf{f}_2(\delta x + \delta s) - \mathbf{m}_{[\omega_{\mathbf{f}_2}]}(\delta n)}{\mathbf{m}_{[\omega_{\mathbf{f}_2}]}(\delta x) - \mathbf{m}_{[\omega_{\mathbf{f}_2}]}(\delta n)}. \tag{39}$$

Then we require

$$\eta \geq \underline{\eta} \tag{40}$$

for acceptance of the step. As a further modification to [LSW17] we increase $[\omega_{\mathbf{f}_2}]$ at least by a fixed factor $\rho_2 \in]1, \rho_1]$ with respect to $[\omega_{\mathbf{f}_2}]^{\text{old}}$, if the decrease condition (40) fails. Moreover, $[\omega_{\mathbf{f}_2}]$ will not be increased, if $\eta \geq \hat{\eta}$ for some $\hat{\eta} \in [\underline{\eta}, 1[$ which is usually chosen close to 1.

4 Local Convergence

In this section the transition of the method to fast local convergence is discussed. Our main point of interest is to show that our flexible choice of retractions will not interfere with fast local convergence of the Lagrange-Newton method. This includes to show that our globalization scheme asymptotically admits full Lagrange-Newton steps.

Throughout this section we impose the following assumptions on the regularity of the problem data and the retractions:

Assumption 4.1. *Let $x_* \in X$ be a local minimizer of f on $c(x) = p$ and $U \subset X$ a neighbourhood of x_*. For $x \in U$ denote $\mathbf{x}_* := (R^X_{x,2})^{-1}x_*$. Bold letters describe the pullbacks to T_xX and T_yY for our given x.*

- *$\mathbf{c}'(0_x)$ is surjective and $\mathbf{L}_1''(0_x, p_x)$ is elliptic on $\ker \mathbf{c}'(0_x)$ with uniform constant $\alpha > 0$ and bounded with uniform constant Γ on $x \in U$.*
- *First order consistent retractions $R^X_{x,i}$ and $R^Y_{c(x),i}$ exist for each $x \in U$ and $i = 1, 2$ and there are constants $\underline{c}, \overline{c} > 0$, such that for all $x, \tilde{x} \in U$:*

  $$\underline{c}\|(R^X_{x_*,2})^{-1}\tilde{x} - 0_{x_*}\| \leq \|(R^X_{x,2})^{-1}\tilde{x} - \mathbf{x}_*\| \leq \overline{c}\|(R^X_{x_*,2})^{-1}\tilde{x} - 0_{x_*}\|. \tag{41}$$

  This is a local norm-equivalence condition on the charts.
- *The assumptions of Proposition 2.1 hold with uniform bounds on the constants $\omega_{\mathbf{c}_2}, \omega_{\mathbf{L}_2}, \omega_{\mathbf{f}_2'}$.*
- *There is a uniform bound γ, such that*

  $$|\mathbf{L}'(0_x, p_x)\Phi''_{X,x}(0_x)(v, w)| \leq \gamma\|\mathbf{x}_* - 0_x\|\|v\|\|w\|. \tag{42}$$

  Taking into account stationarity at x_* this can be seen as a Lipschitz condition on $\mathbf{L}'$, combined with a regularity assumption on Φ_X.

- *There is ω_* independent of x, such that with $\mathbf{x}_* := (R^X_{x,2})^{-1}x_*$:*

$$|\mathbf{f}_2'(\mathbf{x}_*)\mathbf{c}_2'(\mathbf{x}_*)^-(\mathbf{c}_2'(\mathbf{x}_*) - \mathbf{c}_2'(0_x))w| \leq \omega_*\|\mathbf{x}_* - 0_x\|\|w\|.$$

This is a variant of (34).

In the following we consider a sequence x_k, generated by our algorithm. We will show that if x_0 is sufficiently close to x_*, then $x_k \to x_*$ quadratically. Mathematically, taking into account that $(R^X_{x_*,2})^{-1}x_* = 0_{x_*}$ this can be formulated as follows:

$$\exists C_N > 0: \quad \|(R^X_{x_*,2})^{-1}x_{k+1} - 0_{x_*}\| \leq C_N\|(R^X_{x_*,2})^{-1}x_k - 0_{x_*}\|^2. \tag{43}$$

We thus want to observe local quadratic convergence of the iterates, transformed to $T_{x_*}X$ via $R^X_{x_*,2}$.

Lemma 4.1. *Let $x = x_k$, δx the step, computed by our algorithm, and $\mathbf{x}_* := (R^X_{x,2})^{-1}x_*$. Assume that*

$$\exists \tilde{C}_N > 0: \quad \|0_x + \delta x - \mathbf{x}_*\| \leq \tilde{C}_N\|0_x - \mathbf{x}_*\|^2.$$

Then (43) *holds for $x_k = x$ and $x_{k+1} = R^X_{x,2}(0_x + \delta x)$.*

Proof. Let $x_+ = R^X_{x,2}(0_x + \delta x)$ describe one step of our algorithm. Computing

$$\underline{c}\|(R^X_{x_*,2})^{-1}x_+\| \leq \|(R^X_{x,2})^{-1}x_+ - (R^X_{x,2})^{-1}x_*\| = \|0_x + \delta x - \mathrm{x}_*\| \leq \tilde{C}_N\|0_x - \mathbf{x}_*\|^2 \leq \overline{c}^2\tilde{C}_N\|(R^X_{x_*,2})^{-1}x\|^2,$$

yields

$$\|(R^X_{x_*,2})^{-1}x_+\| \leq \tilde{C}_N\frac{\overline{c}^2}{\underline{c}}\|(R^X_{x_*,2})^{-1}x\|^2 = C_N\|(R^X_{x_*,2})^{-1}x\|^2$$

□

In the following we consider full Lagrange-Newton Steps at an iterate $z = (x, p)$

$$\Delta z := (\Delta x, \Delta p) := D_z^2\mathbf{L}_1(0_x, p)^{-1}D_z\mathbf{L}_1(0_x, p), \tag{44}$$

which satisfies the equation

$$\begin{pmatrix} \mathbf{L}_1''(0_x,p) & \mathbf{c}'(0_x)^* \\ \mathbf{c}'(0_x) & 0 \end{pmatrix}\begin{pmatrix} \Delta x \\ \Delta p \end{pmatrix} + \begin{pmatrix} \mathbf{L}'(0_x,p) \\ \mathbf{c}(0_x) \end{pmatrix} = 0.$$

Proposition 4.1. *Suppose that Assumption 4.1 holds close to x_*. Then the full-step variant of our method converges locally quadratically to x_*.*

Proof. We apply one Newton step Δz in T_xX at $z_0 = (0_x, p_x)$ to the following problem:

$$D_z\mathbf{L}(z) = 0 \quad :\Leftrightarrow \quad \begin{pmatrix} \mathbf{L}_2'(v,p) \\ \mathbf{c}_2(v) \end{pmatrix} = 0,$$

which results from a pullback of our original problem to T_xX via $R^X_{x,2}$ and $R^Y_{y,2}$. We obtain with $z_+ = z_0 + \Delta z$ and $z_* = (\mathbf{x}_*, p_*)$:

$$z_+ - z_* = z_0 - z_* + \Delta z = D_z^2\mathbf{L}^{-1}(z_0)(D_z^2\mathbf{L}(z_0)(z_0 - z_*) - D_z\mathbf{L}(z_0))$$

Since we only use norms for the primal component, and p depends on x directly, our aim is to show:

$$\|x_+ - x_*\| \leq C\|0_x - x_*\|^2.$$

Writing the primal component $x_+ - \mathbf{x}_* = n_+ + t_+$ with $\mathbf{c}'(0_x)t_+ = 0$ and $n_+ \perp \ker\mathbf{c}'(0_x)$ we obtain, subtracting $\mathbf{c}_2(\mathbf{x}_*) = 0$:

$$n_+ = \mathbf{c}'(0_x)^-(\mathbf{c}'(0_x)(0_x - \mathbf{x}_*) - (\mathbf{c}(0_x) - \mathbf{c}_2(\mathbf{x}_*))).$$

Application of (34) yields via the fundamental theorem of calculus:

$$\|n_+\| \leq \frac{\omega_{\mathbf{c}_2}}{2}\|0_x - \mathbf{x}_*\|^2 = \frac{\omega_{\mathbf{c}_2}}{2}\|\mathbf{x}_*\|^2.$$

The tangential component t_+ is a minimizer of the problem:

$$\min_{v\in\ker \mathbf{c}'(0_x)} \frac{1}{2}\mathbf{L}_1''(0_x,p_x)(v,v) + (\mathbf{L}_1''(0_x,p_x)(0_x-\mathbf{x}_*) - \mathbf{L}'(0_x,p_x) + \mathbf{L}_1''(0_x,p_x)n_+)v$$

Due to the assumed uniform ellipticity of $\mathbf{L}_1''(0_x,p_x)$ it is sufficient to obtain an estimate for the linear part of this functional of the form:

$$(\mathbf{L}_1''(0_x,p_x)(0_x-\mathbf{x}_*) - \mathbf{L}'(0_x,p_x) + \mathbf{L}_1''(0_x,p_x)n_+)v \leq c\|\mathbf{x}_*\|^2\|v\|. \tag{45}$$

First, we observe that

$$\mathbf{L}_1''(0_x,p_x)(n_+,v) \leq \Gamma\|n_+\|\|v\| \leq \Gamma\frac{\omega_{\mathbf{c}_2}}{2}\|\mathbf{x}_*\|^2\|v\|.$$

Next, we telescope, subtracting $\mathbf{L}_2'(\mathbf{x}_*,p_*)=0$,

$$\begin{aligned}(\mathbf{L}_1''(0_x,p_x)(0_x-\mathbf{x}_*) - \mathbf{L}'(0_x,p_x))v = {}& (\mathbf{L}_1''(0_x,p_x) - \mathbf{L}_2''(0_x,p_x))(0_x-\mathbf{x}_*,v)\\ &+ (\mathbf{L}_2''(0_x,p_x)(0_x-\mathbf{x}_*) - (\mathbf{L}'(0_x,p_x) - \mathbf{L}_2'(\mathbf{x}_*,p_x)))v\\ &+ (\mathbf{L}_2'(\mathbf{x}_*,p_x) - \mathbf{L}_2'(\mathbf{x}_*,p_*))v\end{aligned}$$

into a sum of three terms. The first term is estimated via (22) and (42), taking into account that $v\in\ker\mathbf{c}'(0_x)$:

$$|(\mathbf{L}_1''(0_x,p_x) - \mathbf{L}_2''(0_x,p_x))(0_x-\mathbf{x}_*,v)| = |\mathbf{L}'(0_x,p_x)\Phi_{X,x}''(0_x)(0_x-\mathbf{x}_*,v)| \leq \gamma\|\mathbf{x}_*\|^2\|v\|.$$

The second term is estimated via (35), using the fundamental theorem of calculus:

$$(\mathbf{L}_2''(0_x,p_x)(0_x-\mathbf{x}_*) - (\mathbf{L}'(0_x,p_x) - \mathbf{L}_2'(\mathbf{x}_*,p_x)))v \leq \frac{\omega_{\mathbf{L}_2}}{2}\|\mathbf{x}_*\|^2\|v\|.$$

For the third term, we compute, using $v\in\ker\mathbf{c}'(0_x)$ and $\mathbf{c}'(0_x)\mathbf{c}'(0_x)^- = Id$:

$$(\mathbf{L}_2'(\mathbf{x}_*,p_x) - \mathbf{L}_2'(\mathbf{x}_*,p_*))v = (p_x-p_*)\mathbf{c}'(0_x)\mathbf{c}'(0_x)^-(\mathbf{c}_2'(\mathbf{x}_*) - \mathbf{c}'(0_x))v = (p_x-p_*)\mathbf{c}'(0_x)w.$$

With $w := \mathbf{c}'(0_x)^-(\mathbf{c}_2'(\mathbf{x}_*) - \mathbf{c}'(0_x))v \in \ker\mathbf{c}'(0_x)_\perp$ this yields $\|w\| \leq \omega_{\mathbf{c}_2}\|\mathbf{x}_*\|\|v\|$. We continue, using $p_* = p_*\mathbf{c}_2'(x_*)\mathbf{c}_2'(x_*)^- = -\mathbf{f}_2'(x_*)\mathbf{c}_2'(x_*)^-$:

$$\begin{aligned}|(p_x-p_*)\mathbf{c}'(0_x)w| &= |p_*(\mathbf{c}_2'(\mathbf{x}_*) - \mathbf{c}'(0_x))w - (p_*\mathbf{c}_2'(\mathbf{x}_*) - p_x\mathbf{c}'(0_x))w|\\ &= |-\mathbf{f}_2'(\mathbf{x}_*)\mathbf{c}_2'(\mathbf{x}_*)^-(\mathbf{c}_2'(\mathbf{x}_*) - \mathbf{c}'(0_x))w + (\mathbf{f}_2'(\mathbf{x}_*) - \mathbf{f}'(0_x))w|\\ &\leq \omega_*\|\mathbf{x}_* - 0_x\|\|w\| + \omega_{\mathbf{f}_2}\|\mathbf{x}_* - 0_x\|\|w\| \leq c\|\mathbf{x}_* - 0_x\|^2\|v\|.\end{aligned}$$

Adding all estimates yields (45), as desired. □

Close to an SSC point, we show that the computed, normal and tangential steps, approach to the full Lagrange-Newton steps asymptotically, and from the latter, they inherit local superlinear convergence. On one hand we have that $\delta t = \tau\Delta t$, where $\tau\in(0,1]$ is a damping factor, computed via minimizing

$$\mathbf{m}_{[\omega_{\mathbf{f}_2}]}(\delta x) = \mathbf{f}(0_x) + \mathbf{f}'(0_x)\delta x + \frac{1}{2}\mathbf{L}_1''(0_x,p_x)(\delta x,\delta x) + \frac{[\omega_{\mathbf{f}}]}{6}\|\delta x\|^3 \tag{46}$$

in the affine subspace $\delta n + \operatorname{span}\{\Delta t\}$. We have the relation between the optimization step and the full Lagrange-Newton step Δx:

$$\delta x = \delta n + \delta t = \nu\Delta n + \tau\Delta t, \quad \Delta x = \Delta n + \Delta t.$$

Theorem 4.1. *Assume that x_k converges to the SSC point x_* and assume that the Lipschitz conditions as in Proposition* 2.1 *hold in a neighborhood of x_*. Then we have superlinear convergence.*

Proof. We show that the damping factors ν_k and τ_k tend to 1 as $x \to x_*$. By boundedness of the algorithmic parameter $[\omega_{\mathbf{c}}]$ the normal damping factor ν_k becomes $\nu_k = 1$ eventually, for details see [LSW17].

Concerning τ_k, using the minimizing property of δx_k along the direction Δt_k and by inserting this into the first order optimality conditions for (46), we get that:

$$\begin{aligned}
0 &= \mathbf{m}'_{[\omega_{\mathbf{f}}]}(\delta x_k)\Delta t_k \\
&= (\mathbf{f}'(0_{x_k}) + \mathbf{L}_1''(0_{x_k}, p_{x_k})\delta n_k)\Delta t_k + \mathbf{L}_1''(0_{x_k}, p_{x_k})(\delta t_k, \Delta t_k) + \frac{[\omega_{\mathbf{f}}]}{2}\|\delta x_k\| \langle \delta x_k, \Delta t_k\rangle \\
&= (\mathbf{f}'(0_{x_k}) + \mathbf{L}_1''(0_{x_k}, p_{x_k})\delta n_k)\Delta t_k + \tau_k\Big(\mathbf{L}_1''(0_{x_k}, p_{x_k})(\Delta t_k, \Delta t_k) + \frac{[\omega_{\mathbf{f}}]}{2}\|\delta x_k\| \langle \Delta t_k, \Delta t_k\rangle\Big)
\end{aligned}$$

The equation

$$\begin{aligned}
0 = \mathbf{m}'_0(\delta x_k)\Delta t_k &= (\mathbf{f}'(0_{x_k}) + \mathbf{L}_1''(0_{x_k}, p_{x_k})\delta n_k)\Delta t_k + \mathbf{L}_1''(0_{x_k}, p_{x_k})(\delta t_k, \Delta t_k) \\
&= (\mathbf{f}'(0_{x_k}) + \mathbf{L}_1''(0_{x_k}, p_{x_k})\delta n_k)\Delta t_k + \mathbf{L}_1''(0_{x_k}, p_{x_k}) \langle \Delta t_k, \Delta t_k\rangle
\end{aligned}$$

holds for the full tangential step Δt_k, which minimizes the cubic model $\mathbf{m}_{[\omega_{\mathbf{f}}]}$ for $[\omega_{\mathbf{f}}] = 0$. Subtracting these two equations, we obtain

$$\mathbf{L}_1''(0_{x_k}, p_{x_k})(\Delta t_k, \Delta t_k) = \tau_k \left[\mathbf{L}_1''(0_{x_k}, p_{x_k})(\Delta t_k, \Delta t_k) + \frac{[\omega_{\mathbf{f}}]}{2}\|\delta x_k\| \langle \Delta t_k, \Delta t_k\rangle\right] \tag{47}$$

then

$$\tau_k = \frac{\mathbf{L}_1''(0_{x_k}, p_{x_k})(\Delta t_k, \Delta t_k)}{\mathbf{L}_1''(0_{x_k}, p_{x_k})(\Delta t_k, \Delta t_k) + \frac{[\omega_{\mathbf{f}}]}{2}\|\delta x_k\| \langle \Delta t_k, \Delta t_k\rangle} \tag{48}$$

With that we perform the following estimate, which holds sufficiently close to x_*:

$$\|0_{x_k} - \mathbf{x}_*\| \le \tilde{C}\|\Delta x_k\| \le \frac{\tilde{C}}{\tau_k}\|\delta x_k\| \le \tilde{C}\left(1 + \frac{[\omega_{\mathbf{f}_2}]}{2\alpha}\|\delta x_k\|\right)\|\delta x_k\| \le C(1 + [\omega_{\mathbf{f}_2}]\|\delta x_k\|)\|\delta x_k\|, \tag{49}$$

where α is the ellipticity constant of $\mathbf{L}_1''$.

Next, consider the acceptance test (39). Since $\mathbf{m}_{[\omega_{\mathbf{f}_2}]}(\delta x_k) < \mathbf{m}_{[\omega_{\mathbf{f}_2}]}(\delta n_k)$, (39) is certainly fulfilled with $\eta \ge 1$, if $\mathbf{f}(\delta x_k + \delta s_k) \le \mathbf{m}_{[\omega_{\mathbf{f}_2}]}(\delta x_k)$. To establish such an estimate, we compute from Proposition 2.1 and (49).

$$\mathbf{f}(\delta x_k + \delta s_k) - \mathbf{q}(\delta x_k) \le C\|\delta x_k\|^3 + C\gamma\|0_{x_k} - \mathbf{x}_*\|\|\delta x_k\|^2 \le C(1 + [\omega_{\mathbf{f}_2}]\|\delta x_k\|)\|\delta x_k\|^3$$

Since

$$\mathbf{m}_{[\omega_{\mathbf{f}_2}]} - \mathbf{q}(\delta x_k) = \frac{[\omega_{\mathbf{f}_2}]}{6}\|\delta x_k\|^3$$

we obtain $f(\delta x_k + \delta s_k) \le \mathrm{m}_{[\omega_{\mathbf{f}_2}]}(\delta x_k)$, if

$$6C(1 + [\omega_{\mathbf{f}_2}]\|\delta x_k\|) \le \frac{[\omega_{\mathbf{f}_2}]}{6}$$

For sufficiently small δx_k this is true, if

$$[\omega_{\mathbf{f}_2}] \ge \frac{6C}{1 - 6C\|\delta x_k\|}.$$

Thus, we conclude that close to a minimizer (39) always holds with $\eta \ge 1 > \hat{\eta}$, if $[\omega_{\mathbf{f}_2}]$ is above a certain bound that only depends on the problem and the chosen neighbourhood around x_*. Consequently, by our algorithmic mechanism, $[\omega_{\mathbf{f}_2}]$ cannot become unbounded.

Hence, as $x_k \to x_*$, implies by (48) that $\tau_k \to 1$ taking ellipticity of $\mathbf{L}_1''$ close to x_* into account. Thus, we obtain local superlinear convergence of our algorithm. More accurrately, by boundedness of $[\omega_{\mathbf{f}_2}]$ we obtain, using $\|\delta x_k\| \leq \|\Delta x_k\|$ and (48):

$$\tau_k \geq \frac{1}{1 + C\|\Delta x_k\|} \quad \Rightarrow \quad 1 - \tau_k \leq C\|\Delta x_k\|$$

and hence

$$\|\Delta x_k - \delta x_k\| \leq (1 - \tau_k)\|\Delta x_k\| \leq C\|\Delta x_k\|^2.$$

Since $\|\delta s_k\| \leq C\|\Delta x_k\|^2$ as well, we have

$$\|\Delta x_k - (\delta x_k + \delta s_k)\| \leq C\|\Delta x_k\|^2,$$

so quadratic convergence of the full Newton method carries over to our globalized version. □

5 Application: Inextensible Flexible Rods

In this section we consider the numerical simulation of flexible inextensible rods to illustrate our approach. Flexible rods are present in many real life problems, for example engineers are interested in the static and dynamic behaviour of flexible pipelines used in off-shore oil production under the effects of streams, waves, and obstacles; or protein structure comparison [LSZ11] where elastic elastic curves are used to represent and compare protein structures. Here we consider the problem where the stable equilibrium position of an inextensible transversely isotropic elastic rod under dead load is searched. First we provide the formulation and the mathematical analysis of the problem, followed by the discretization and the derivatives of the mappings over the manifold of kinematically admissible configurations. We finish with some numerical experiments.

5.1 Problem formulation

Here we provide the energetic formulation of the problem of finding the stable equilibrium position of an inextensible, transversely isotropic elastic rod under dead loading. For more details on the derivation of the model see [GLT89]. We consider the following minimization problem

$$\min_{y \in V} J(y) \tag{50}$$

where the energy J and the manifold V which describes the inextensibility condition are given by:

$$J(y) = \frac{1}{2}\int_0^1 EI\,\langle y'', y''\rangle\,ds - \int_0^1 \langle g, y\rangle\,ds, \tag{51}$$

$$V = \{y \mid y \in H^2([0,1];\mathbb{R}^3), |y'(s)| = 1 \text{ on } [0,1]\}. \tag{52}$$

with boundary conditions

$$\begin{aligned} y(0) &= y_a \in \mathbb{R}^3, \;\; y'(0) = y_a' \in \mathbb{S}^2 \\ y(1) &= y_b \in \mathbb{R}^3, \;\; y'(1) = y_b' \in \mathbb{S}^2 \end{aligned} \tag{53}$$

Above $EI(s) > 0$ is the flexural stiffness of the rod, g is the lineic density of external loads, and y', y'' are the the derivatives of y with respect to $s \in [0,1]$ and $\mathbb{S}^2$ is the unit sphere

$$\mathbb{S}^2 = \{v \in \mathbb{R}^3 : |v| = 1\}.$$

We reformulate (50) as:

$$\min_{(y,v) \in Y \times V} f(y,v) \;\; s.t \;\; y' - v = 0. \tag{54}$$

with

$$f(y,v) = \frac{1}{2}\int_0^1 EI\,\langle v', v'\rangle\,ds - \int_0^1 \langle g, y\rangle\,ds, \tag{55}$$

$$\begin{aligned} Y &= \{y \in H^2([0,1];\mathbb{R}^3) \ : \ y(0) = y_a,\ y(1) = y_b\} \\ V &= \{v \in H^1([0,1];\mathbb{S}^2) \ : \ v(0) = v_a,\ v(1) = v_b\} \end{aligned} \tag{56}$$

From the formulation given in (54) we get the constrained minimization problem:

$$\min_{(y,v)\in(Y\times V)} f(y,v) \ \ s.t \ \ c(y,v) = 0$$

where Y and V are given by:

$$\begin{aligned} Y &= H^2([0,1];\mathbb{R}^3), \\ V &= H^1([0,1];\mathbb{S}^2). \end{aligned}$$

5.2 Mathematical analysis of the problem

Concerning the study of existence and the uniqueness of the solutions of the problem (50) we refer the reader to the books [GLT89, AR78] for a detailed and complete mathematical analysis of these kind of problems. In the following we assume that $EI \in L^\infty([0,1])$ is non-negative. Concerning to the existence properties of the problem (50) we have the following theorem.

Theorem 5.1. *Suppose that $|y_a - y_b| < 1$, (53) holds, and that the linear functional $y \to \int_0^1 \langle g, y\rangle ds$ is continuous on $H^2([0,1];\mathbb{R}^3)$. Then the problem (50) has at least one solution.*

Proof. See [GLT89]. □

First order optimality conditions Here we derive the first order optimality conditions and the corresponding KKT-system for the problem as formulated in (54), namely:

$$\begin{aligned} \min_{(y,v)\in Y\times V} \frac{1}{2}\int_0^1 EI\,\langle v', v'\rangle\,ds - \int_0^1 \langle g, y\rangle\,ds, \\ s.t \qquad y' - v = 0, \\ y(0) = y_a \in \mathbb{R}^3,\ y(1) = y_b \in \mathbb{R}^3 \\ v(0) = v_a \in \mathbb{S}^2,\ v(1) = v_b \in \mathbb{S}^2. \end{aligned}$$

Defining the Lagrangian function

$$L(y,v,p) = f(y,v) + pc(y,v) = \frac{1}{2}\int_0^1 EI\,\langle v', v'\rangle - \langle g, y\rangle + \langle p, y' - v\rangle\,ds$$

for $p \in P = L_2([0,1],\mathbb{R}^3)$ we obtain:

$$\begin{aligned} L_y(y,v,p)\delta y &= \int_0^1 -\langle g, \delta y\rangle + \langle p, \delta y'\rangle\ ds, \\ L_v(y,v,p)\delta v &= \int_0^1 EI\,\langle v', \delta v'\rangle - \langle p, \delta v\rangle\,ds \end{aligned}$$

yielding the KKT-sytem

$$\begin{aligned} \int_0^1 -\langle g, \delta y\rangle + \langle p, \delta y'\rangle\,ds = 0 \quad \forall \delta y \in Y \\ \int_0^1 EI\,\langle v', \delta v'\rangle - \langle p, \delta v\rangle\,ds = 0 \quad \forall \delta v \in T_v V \\ \int_0^1 \langle \delta p, y' - v\rangle\,ds = 0 \quad \forall \delta p \in P. \end{aligned}$$

5.3 Finite difference approximation of the problem

For discretization, we use a very simple finite difference approach. We discretize the interval $[0,1]$ uniformly

$$s_i = ih, \;\; i = 0, ..., n-1$$

where $h = \frac{1}{n-1}$. Evaluating at each nodal point we denote

$$\begin{aligned} y(s_i) &= y_i \in \mathbb{R}^3, \;\; i = 0, ..., n-1 \\ v(s_i) &= v_i \in \mathbb{S}^2, \;\; i = 0, ..., n-1. \end{aligned} \tag{57}$$

with boundary conditions:

$$\begin{aligned} y(0) &= y_a \in \mathbb{R}^3, \qquad & y(1) &= y_b \in \mathbb{R}^3 \\ v(0) &= v_a \in \mathbb{S}^2, \qquad & v(1) &= v_b \in \mathbb{S}^2. \end{aligned}$$

Employing forward finite difference discretization and a Riemann sum for the integrals yields the following approximation of the energy functional

$$f(y,v) = \frac{1}{2} \sum_{i=0}^{n-1} h \left\langle \frac{1}{h}(v_{i+1} - v_i), \frac{1}{h}(v_{i+1} - v_i) \right\rangle - \sum_{i=1}^{n} h \left\langle g_i, y_i \right\rangle . \tag{58}$$

Concerning the constraint $c(y,v)$, performing forward finite differences to the equation $y' - v = 0$, the discretized constraint mapping takes the form

$$\frac{y_{i+1} - y_i}{h} - v_i = 0 \quad i = 0, ..., n-1. \tag{59}$$

We observe that the codomain of our constraint mapping is a linear space, which eliminates the need for a retraction in the codomain.

In the formulation above of the discrete inextensible rod, the manifold X is $(\mathbf{R}^3 \times \mathbb{S}^2)^n$, with n the number of grid vertices. The elements of the manifold X are denoted by the cartesian product

$$(y,v) = \prod_{i=0}^{n-1} (y_i, v_i), \qquad y_i \in \mathbb{R}^3, \; v_i \in \mathbb{S}^2.$$

The tangent space at $(y,v) \in X$ is given by the following direct sum of vector spaces

$$T_{(y,v)}X = \bigoplus_{i=0}^{n-1} \left(T_{y_i}\mathbb{R}^3 \oplus T_{v_i}\mathbb{S}^2 \right).$$

The update, using the retraction map $R_{(y,v)}T_{(y,v)}X \longrightarrow X$, is done in a component-wise way by:

$$(y^+, v^+) = R_{(y,v)}(\delta y, \delta v) = \prod_{i=0}^{n-1} (y_i + \delta y_i, R_{v_i}(\delta v_i)).$$

5.4 Retractions and their implementation via local parametrizations

As presented above, retractions are defined as mappings $R_x^M : T_xM \to M$. For their implementation on a computer, we have to choose a basis of T_xM and represent R_x^M with respect to that basis. This yields the concept of local parametrizations as described in [AMS09]. A local parametrization is a map $\mu_x : \mathbb{R}^d \longrightarrow M \;\; \mu_x(0) = x$, that is a local diffeomorphism around $\mathbb{R}^d$. Parametrizations can be defined from retractions by selecting a basis $\{\xi_1, ..., \xi_d\}$ of T_xM and defining

$$\mu_x(u_1, ..., u_d) = R_x(u_1\xi_1 + .. + u_d\xi_d).$$

For our case we construct local parametrizations around each node v_i on each sphere, induced by retractions $R^{\mathbb{S}^2}$, this is, we look for local diffeomorphisms around $v_i \in \mathbb{S}^2$:

$$\begin{aligned} \mu_{v_i} : \mathbb{R}^2 &\longrightarrow \mathbb{S}^2 \\ u &\longrightarrow \mu_{v_i}(u) \end{aligned}$$

such that

$$\mu_{v_i}(0) = v_i \ \ and \ \ \mu_{v_i}(u) \in \mathbb{S}^2.$$

In the following we will consider two alternative retractions, implemented via suitable local parametrizations:

Projection to the sphere. In the following we use the representation:

$$T_v\mathbb{S}^2 = \{w \in \mathbb{R}^3 : w \perp v\}.$$

For $v \in \mathbb{S}^2$, let be $u \in \mathbb{R}^2$, $u = (u_1, u_2)$ and $\{\zeta_1, \zeta_2\} \in T_v\mathbb{S}^2$ be an orthogonal basis for the tangent space of $\mathbb{S}^2$ at every v. We define the parametrization around v by:

$$\mu_{v,p}(u) = \frac{v + u_1\zeta_1 + u_2\zeta_2}{\|v + u_1\zeta_1 + u_2\zeta_2\|}.$$

This parametrization implements the retraction:

$$R_{v,p}(\delta v) = \frac{v + \delta v}{\|v + \delta v\|}$$

and they satisfy:

$$\begin{aligned} R_{v,p}(0) = \mu_{v,p}(0) &= v, \\ DR_{v,p}(0) &= id_{T_v\mathbb{S}^2}. \end{aligned}$$

Details can be found in [AMS09].

Proposition 5.1. *Let be $v \in \mathbb{S}^2$ and suppose that $\{\zeta_1, \zeta_2\}$ is an orthonormal basis of the tangent space to $\mathbb{S}^2$ at v and let be $\mu_{v,p}$ the parametrization around $v \in \mathbb{S}^2$ given by:*

$$\mu_{v,p}(u) = \frac{v + u_1\zeta_1 + u_2\zeta_2}{\|v + u_1\zeta_1 + u_2\zeta_2\|}.$$

Then we have that

i)

$$\mu'_{v,p}(0)\delta u = \delta u_1\zeta_1 + \delta u_2\zeta_2$$

ii)

$$\mu''_{v,p}(0)(\delta u, \delta w) = -(\delta u_1\delta w_1 + \delta u_2\delta w_2)v.$$

Matrix exponential. The following alternative retraction uses a characterisation of $T_v\mathbb{S}^2$ via the space of skew-symmetric matrices $\mathfrak{so}(3) = \{H \in \mathbb{R}^{3\times3} | H = -H^T\}$:

$$T_v\mathbb{S}^2 = \{Hv : H \in \mathfrak{so}(3)\}.$$

This follows from $\langle Hv, v\rangle = -\langle v, Hv\rangle = 0$ by the fact that Hv can be written as $w \times v$, which is non-zero if $0 \neq w \perp v$.

Using the matrix exponential map, and setting $\delta v = Hv$ we can define the following retraction:

$$R_{v,e}(\delta v) = \exp(H)v.$$

where

$$\exp : \mathfrak{so}(3) \longrightarrow SO(3)$$

is the matrix exponential mapping with $SO(3) = \{Q \in \mathbb{R}^{3\times 3} | QQ^T = Q^TQ = Id_3, \det(Q) = 1\}$, the group of rotations. We remark that the retraction is well defined: if $H_0 v = 0$, then $\exp(H_0)v = v$ as can be seen by the series expansion of the matrix exponential, and thus $\exp(H + H_0)v = \exp(H)v$.

For any given $v \in \mathbb{S}^2$ we consider the following basis for the tangent space $T_v\mathbb{S}^2$

$$b_2 = \{C_1 v, C_2 v\} \tag{60}$$

where $C_j \in \mathfrak{so}(3)$ are chosen in a way that $C_j v \neq 0$ for $j = 1, 2$. Now we define the map for $u = (u_1, u_2)$:

$$\mu_{v,e}(u) = \exp\left(u_1 C_1 + u_2 C_2\right) v$$

Since $\exp\left(u_1 C_1 + u_2 C_2\right) \in SO(3)$ and $\exp(0) = I$ we obtain:

$$\langle \mu_{v,e}(u), \mu_{v,e}(u) \rangle = 1 \ \textit{and} \ \mu_{v,e}(0) = v.$$

which means that $\mu_{v,e}(u) \in \mathbb{S}^2$.

Proposition 5.2. *Let be $v \in \mathbb{S}^2$ and C_1 and C_2 as defined above. Consider the parametrization around $v \in \mathbb{S}^2$ given by:*

$$\mu_{v,e}(u) = \exp\left(u_1 C_1 + u_2 C_2\right) v.$$

Then we have that:

$$\mu'_{v,e}(0)\delta u = (\delta u_1 C_1 + \delta u_2 C_2)v.$$

consequently, derivative of the retraction reads:

$$DR_{v,e}(0) = id_{T_v\mathbb{S}^2}.$$

Additionally, we have that

$$\mu''_{v,e}(0)(\delta u, \delta w) = (\delta u_1 C_1 + \delta u_2 C_2)(\delta w_1 C_1 + \delta w_2 C_2)v$$

Proposition 5.3. *The parametrizations $\mu_{v,p}$ and $\mu_{v,e}$ induce retractions $R_{v,p}$ and $R_{v,e}$ that are first order consistent.*

Proof. From Lemma 2.1 we now that every pair of first order retractions are first order consistent and from propositions 5.1 and 5.2 we have that $\mu_{v,p}$ and $\mu_{v,e}$ are of first order, therefore they are first order consistent. □

Remark 5.1. *We stress that the choice of basis of T_xM that has to be made for the definition of the parameterization μ_v does not affect the definition of the retraction R_v. This decouples the representation of the steps from the representation of the iterates.*

5.5 The pullback of the discretized problem

We now pullback the energy functional f and the constraint mapping c using a local parametrization at each v_i through μ_{v_i}, which denotes any of the two parametrizations, presented above. From (58) the pullbacked energy functional takes the form:

$$\mathbf{f}(y, u) = \frac{EI}{2} \sum_{i=0}^{n-1} h \left\langle \frac{1}{h}(\mu_{v_{i+1}}(u_{i+1}) - \mu_{v_i}(u_i)), \frac{1}{h}(\mu_{v_{i+1}}(u_{i+1}) - \mu_{v_i}(u_i)) \right\rangle - \sum_{i=0}^{n-1} h \left\langle g_i, y_i \right\rangle. \tag{61}$$

and

$$\mathbf{c}_i(y,u) = \frac{y_{i+1}-y_i}{h} - \mu_{v_i}(u_i) = 0 \tag{62}$$

for $i = 0, ..., n-1$, and where

$$\mu_{v_i}(u_i) : \mathbb{R}^2 \longrightarrow \mathbb{S}^2$$

is a local parametrization around v_i.

Derivatives of the pullbacked quantities. Now we provide the derivatives of the involved pullbacked mappings. This is done through the composition with the local parametrizations of the sphere. The derivatives are computed centered at the zero of each tangent space parametrization of each sphere $\mathbb{S}^2$. For the composite step method, we need to compute first and second derivatives of both, energy and constraint mappings in charts.

Proposition 5.4. *Consider the discretized energy functional in* (61)*. Its first and second derivatives are given by:*

$$\boldsymbol{f}'(y,u) = \frac{\partial \boldsymbol{f}(y,u)}{\partial y}\delta y + \frac{\partial \boldsymbol{f}(y,u)}{\partial u}\delta u$$

and

$$\boldsymbol{f}''(y,u) = \begin{bmatrix} \delta y & \delta u \end{bmatrix} \begin{bmatrix} \frac{\partial^2 \boldsymbol{f}(y,u)}{\partial y^2} & 0 \\ 0 & \frac{\partial^2 \boldsymbol{f}(y,u)}{\partial u^2} \end{bmatrix} \begin{bmatrix} \delta y \\ \delta u \end{bmatrix}$$

where

$$\frac{\partial \boldsymbol{f}(y,u)}{\partial y_i}\delta y = -h\langle f_i, \delta y\rangle, \qquad \frac{\partial^2 \boldsymbol{f}(y,u)}{\partial y_i^2} = 0,$$

and, at $u = 0$, taking into account that $\mu_{v_i}(0) = v_i$:

$$\begin{aligned}
\frac{\partial \boldsymbol{f}(y,0)}{\partial u_i}\delta u &= \frac{1}{h}\left\langle \mu'_{v_i}(0)\delta u, v_i - v_{i-1}\right\rangle - \frac{1}{h}\left\langle \mu'_{v_i}(0)\delta u, v_{i+1} - v_i\right\rangle \\
&= -\frac{1}{h}\left\langle \mu'_{v_i}(0)\delta u, v_{i+1} - 2v_i + v_{i-1}\right\rangle \\
\frac{\partial^2 \boldsymbol{f}(y,0)}{\partial u_i^2}(\delta u, \delta w) &= -\frac{1}{h}\left\langle \mu''_{v_i}(0)(\delta u, \delta w), v_{i+1} - 2v_i + v_{i-1}\right\rangle + \frac{2}{h}\left\langle \mu'_{v_i}(0)\delta u, \mu'_{v_i}(0)\delta w\right\rangle \\
\frac{\partial^2 \boldsymbol{f}(y,0)}{\partial u_i \partial u_{i-1}}(\delta u, \delta w) &= -\frac{1}{h}\left\langle \mu'_{v_i}(0)\delta u, \mu'_{v_{i-1}}(0)\delta w\right\rangle.
\end{aligned}$$

Proposition 5.5. *The discretized constraint mapping $\boldsymbol{c}$ in* (62) *has the following derivatives:*

$$\begin{aligned}
\boldsymbol{c}'_i(y,0)(\delta y, \delta u) &= -\frac{1}{h}\delta y_i - \mu'_{v_i}(0)\delta u_i \\
\boldsymbol{c}''_i(y,0)(\delta y, \delta u)^2 &= \begin{bmatrix} \delta y_i & \delta u_i \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & -\mu''_{v_i}(0) \end{bmatrix} \begin{bmatrix} \delta y_i \\ \delta u_i \end{bmatrix}.
\end{aligned}$$

6 Numerical Results

We provide numerical simulations in order to ilustrate the performance of the composite step method. We remind the problem setting:

$$\min_{(y,v)\in Y\times V} \frac{1}{2}\int_0^1 EI\,\langle v', v'\rangle\, ds - \int_0^1 \langle g, y\rangle\, ds \quad s.t. \quad y' - v = 0$$

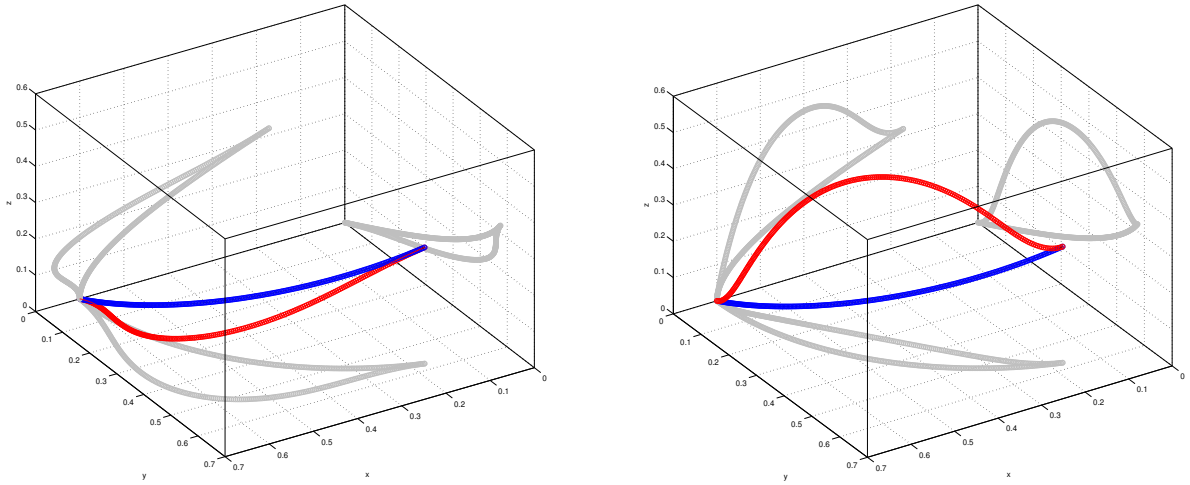

Figure 1: Solutions of the rod problem, blue initial configuration, red computed solution, grey shades: projection to the coordinate planes. Left: rod without external force. Right: rod with external force $1000\, e_3$, (pointing upwards)

where $EI > 0$ is the stiffness of the rod, and g describe the external loads. As initial configuration we consider a rod (y_0, v_0) which assumes the form:

$$y_0(s) = [r\cos(\omega s), r\sin(\omega s), a^2\omega s], \qquad v_0(s) = y_0'(s)$$

with $s \in [0,1]$ $r > 0$, $a > 0$ and $\omega = \frac{1}{\sqrt{r^2+a^2}}$. The rod is clamped at $y_a = y_0(0) = [r,\, 0,\, 0]$ $y_b = y_0(1) = [r\cos(\omega),\, r\sin(\omega),\, a^2\omega]$ and $v_a = y_0'(0)/|y_0'(0)|$, $v_b = y_0'(1)/|y_0'(1)|$. We perform numerical simulations for $r = 0.6$, $a = 0.5$. The stiffness of the rod will be constant and given by $EI = 1.0$. A minimization without external forces, using the exponential retraction μ_2 and $n = 240$ nodes converges in 7 iterations. The corresponding result can be seen in Figure 1, left.

$R_1 \setminus R_2$	$\mu_{v,p}$	$\mu_{v,e}$
$\mu_{v,p}$	9	9
$\mu_{v,e}$	10	10

Table 1: Number of composite step iterations for different combinations of retraction. The pullback is done with the parametrization in the column and the update with the parametrization in the row. Here $\mu_{v,p}(\xi) = \frac{v+\xi}{\|v+\xi\|}$ and $\mu_{v,e}(\xi) = \exp(\xi)v$.

n	#iterations
120	9
240	12
480	8
960	10

Table 2: Number of composite step iterations for the problem with different number of nodes n. The pullback and updates are done with the parametrization $\mu_{v,e}(\xi) = \exp(\xi)v$.

Next, we apply an external force $g = 1000e_3$ to the rod, where $e_3 = [0,0,1]^T$ (cf. Figure 1, right). We consider the two dicussed retractions and combinations of them and observe similar numbers of iterations in all cases (cf. Table 1). Also the number of iterations is largely independent of the size of the grid (cf. Table 2).

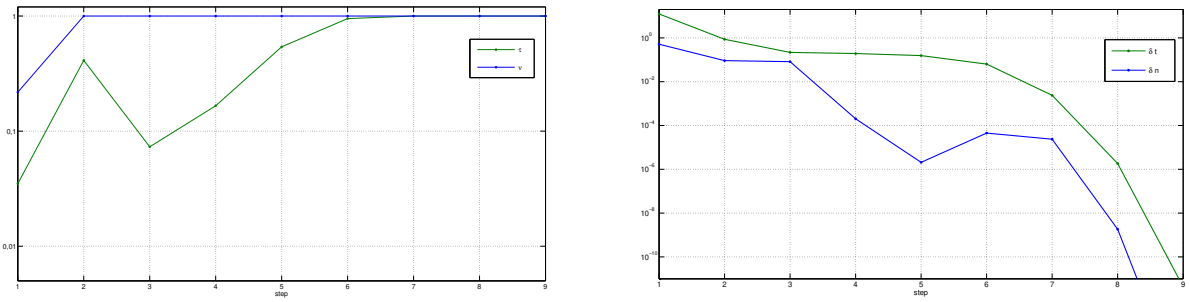

Figure 2: Iteration history: left: damping factors for normal and tangent steps, right: lengths of steps.

In Figure 2 we take a closer look at the iteration history. We observe that after the globalization phase the damping factors are 1 eventually and that the step sizes δt and δn become very small, close to the solution, indicating local superlinear convergence.

Acknowledgements

This work was supported by the DFG grant SCHI 1379/3-1 "Optimierung auf Mannigfaltigkeiten für die numerische Lösung von gleichungsbeschänkten Variationsproblemen"

References

[AKT12] François Alouges, Evaggelos Kritsikis, and Jean-Christophe Toussaint. A convergent finite element approximation for landau–lifschitz–gilbert equation. *Physica B: Condensed Matter*, 407(9):1345–1349, 2012.

[Alo97] François Alouges. A new algorithm for computing liquid crystal stable configurations: the harmonic mapping case. *SIAM journal on numerical analysis*, 34(5):1708–1726, 1997.

[AMS09] Pierre-Antoine Absil, Robert Mahony, and Rodolphe Sepulchre. *Optimization algorithms on matrix manifolds*. Princeton University Press, 2009.

[AR78] Stuart S Antman and Gerald Rosenfeld. Global behavior of buckled states of nonlinearly elastic rods. *Siam Review*, 20(3):513–566, 1978.

[Bal02] John M Ball. Some open problems in elasticity. In *Geometry, mechanics, and dynamics*, pages 3–59. Springer, 2002.

[BP07] Sören Bartels and Andreas Prohl. Constraint preserving implicit finite element discretization of harmonic map flow into spheres. *Mathematics of Computation*, 76(260):1847–1859, 2007.

[BS89] J Badur and Helmut Stumpf. *On the influence of E. and F. Cosserat on modern continuum mechanics and field theory*. Ruhr-Universität Bochum, Institut für Mechanik, 1989.

[CGT00] Andrew R. Conn, Nicholas I.M. Gould, and Philippe L. Toint. *Trust region methods*, volume 1. Siam, 2000.

[Deu11] Peter Deuflhard. *Newton methods for nonlinear problems: affine invariance and adaptive algorithms*, volume 35. Springer Science & Business Media, 2011.

[EL78] James Eells and Luc Lemaire. A report on harmonic maps. *Bulletin of the London mathematical society*, 10(1):1–68, 1978.

[GLT89] Ronald Glowinski and Patrick Le Tallec. *Augmented Lagrangian and operator-splitting methods in nonlinear mechanics*, volume 9. SIAM, 1989.

[HT04] Knut Huper and Jochen Trumpf. Newton-like methods for numerical optimization on manifolds. In *Signals, Systems and Computers, 2004. Conference Record of the Thirty-Eighth Asilomar Conference on*, volume 1, pages 136–139. IEEE, 2004.

[KVBP+14] Evaggelos Kritsikis, A Vaysset, LD Buda-Prejbeanu, François Alouges, and J-C Toussaint. Beyond first-order finite element schemes in micromagnetics. *Journal of Computational Physics*, 256:357–366, 2014.

[LL89] San-Yih Lin and Mitchell Luskin. Relaxation methods for liquid crystal problems. *SIAM Journal on Numerical Analysis*, 26(6):1310–1324, 1989.

[LSW14] Lars Lubkoll, Anton Schiela, and Martin Weiser. An optimal control problem in polyconvex hyperelasticity. *SIAM Journal on Control and Optimization*, 52(3):1403–1422, 2014.

[LSW17] Lars Lubkoll, Anton Schiela, and Martin Weiser. An affine covariant composite step method for optimization with pdes as equality constraints. *Optimization Methods and Software*, 32(5):1132–1161, 2017.

[LSZ11] Wei Liu, Anuj Srivastava, and Jinfeng Zhang. A mathematical framework for protein structure comparison. *PLoS Computational Biology*, 7(2):e1001075, 2011.

[Lue72] David G Luenberger. The gradient projection method along geodesics. *Management Science*, 18(11):620–631, 1972.

[MB11] Peter W. Michor Martin Bauer, Philipp Harms. Sobolev metrics on shape space of surfaces. *Journal of Geometric Mechanics*, 3(1941 4889 2011 4 389):389, 2011.

[Mie02] Alexander Mielke. Finite elastoplasticity lie groups and geodesics on sl (d). In *Geometry, mechanics, and dynamics*, pages 61–90. Springer, 2002.

[MS04] Nicholas Manton and Paul Sutcliffe. *Topological solitons*. Cambridge University Press, 2004.

[PFA06] Xavier Pennec, Pierre Fillard, and Nicholas Ayache. A Riemannian framework for tensor computing. *International Journal of computer vision*, 66(1):41–66, 2006.

[Pro95] Jacques Prost. *The physics of liquid crystals*, volume 83. Oxford university press, 1995.

[RW12] Wolfgang Ring and Benedikt Wirth. Optimization methods on Riemannian manifolds and their application to shape space. *SIAM Journal on Optimization*, 22(2):596–627, 2012.

[Sch14] Volker Schulz. A Riemannian view on shape optimization. *Foundations of Computational Mathematics*, 14(3):483–501, 2014.

[SS00] Jalal M Ihsan Shatah and Michael Struwe. *Geometric wave equations*, volume 2. American Mathematical Soc., 2000.

[SSW15] Volker Schulz, Martin Siebenborn, and Kathrin Welker. Towards a Lagrange–Newton approach for pde constrained shape optimization. In *New Trends in Shape Optimization*, pages 229–249. Springer, 2015.

[TSC00] Bei Tang, Guillermo Sapiro, and Vicent Caselles. Diffusion of general data on non-flat manifolds via harmonic maps theory: The direction diffusion case. *International Journal of Computer Vision*, 36(2):149–161, 2000.